O Português Afro-Brasileiro

O Português Afro-Brasileiro

Dante Lucchesi
Alan Baxter
Ilza Ribeiro
(Organizadores)

EDUFBA
2009

Projeto gráfico, capa e editoração eletrônica
Alana Gonçalves de Carvalho

Arte final de capa
Joe Lopes

Editoração eletrônica
Heloisa O. de S. e Castro

Revisão
Vera Rollemberg

Imagem da capa
Seu José e Dona Rita, casal de idosos da comunidade de Cinzento, em foto de Sabiá, gentilmente cedida pelo filho do casal, Seu Saviano, e sua esposa, Dona Delci.

Sistema de Bibliotecas - UFBA

O português afro-brasileiro / Dante Lucchesi, Alan Baxter, Ilza Ribeiro (Organizadores). - Salvador : EDUFBA, 2009.
576 p. il.

ISBN 978-85-232-0596-6

1. Língua portuguesa - Brasil - Africanismos. 2. Línguas africanas. 3. Lingüística histórica - Brasil. 4. Sociolingüística - Brasil. 5. Dialetos crioulos - Bahia. I. Lucchesi, Dante. II. Baxter, Alan. III. Ribeiro, Ilza.

CDD - 469.798

EDUFBA
Rua Barão de Jeremoabo, s/n, *Campus* de Ondina,
40170-115, Salvador-BA, Brasil
Tel/fax: (71) 3283-6164
www.edufba.ufba.br | edufba@ufba.br

Aos afro-brasileiros

Sumário

Apresentação **15**

Os autores **23**

Introdução **27**
Dante Lucchesi

1ª Parte
O contexto sócio-histórico e os fundamentos teóricos e metodológicos

1 História do contato entre línguas no Brasil **41**
Dante Lucchesi

1.1 Panorama da história sociolinguística do contato entre línguas no Brasil 43
1.2 O contato do português com as línguas africanas 57
1.2.1 A proporção de africanos na formação da sociedade brasileira 59
1.2.2 Fatores que determinaram a ausência de um processo representativo de crioulização do português no Brasil 62
1.3 A transmissão linguística irregular na formação do português brasileiro 71

2 O português afro-brasileiro: as comunidades analisadas **75**
Dante Lucchesi, Alan Baxter, Jorge Augusto Alves da Silva, Cristina Figueiredo

2.1 As comunidades de fala analisadas 83
2.2 A comunidade de Helvécia 85
2.2.1 Caracterização socioeconômica e demográfica 86
2.2.2 Sócio-história 87
2.2.3 Caracterização linguística feita no início da década de 1960 91
2.2.4 Conclusão: o diagnóstico do caso Helvécia e suas implicações para a história sociolinguística do Brasil 92
2.3 A comunidade de Cinzento 95
2.4 Os arraiais de Rio de Contas 97
2.5 A comunidade de Sapé 98
2.6 Conclusão 100

3 A transmissão linguística irregular **101**
Dante Lucchesi, Alan Baxter

3.1 A formação de variedades linguísticas em situação de contato como um *continuum* de erosão e reestruturação gramatical 102
3.2 Parâmetros sócio-históricos 107
3.3 Aspectos linguísticos 109
3.4 Mecanismos de expansão linguística envolvidos na transmissão irregular 114
3.5 A natureza crioulizante de uma variedade linguística 118
3.6 A transmissão linguística irregular como um processo variável 121

4 Teorias da estrutura e da mudança linguísticas e o contato entre línguas **125**
Dante Lucchesi, Ilza Ribeiro

4.1 A teoria da gramática e o estudo dos padrões coletivos de comportamento linguístico 126
4.1.1 Estudo dos padrões coletivos de comportamento linguístico 128
4.1.2 A mudança linguística e a teoria da gramática gerativa 131
4.1.3 Síntese dos desenvolvimentos recentes e as alternativas que se colocam para a conjunção dos dois modelos 140
4.2 A competição entre gramáticas nas situações de contato entre línguas 144
4.2.1 A concorrência entre gramáticas na história sociolinguística do Brasil 147
4.2.2 O princípio da coesão estrutural 150

5 A metodologia **155**
Dante Lucchesi

5.1 O acervo de fala vernácula do português afro-brasileiro 155
5.1.1 A estrutura das amostras de fala vernácula 157
5.1.2 O caráter das entrevistas 159
5.1.3 Os temas das entrevistas 160
5.1.4 O tratamento do áudio 162
5.1.5 A transcrição 162
5.2 O suporte quantitativo 163

2ª Parte
A gramática do português afro-brasileiro

6 A realização do sujeito pronominal **167**
Dante Lucchesi

6.1 O parâmetro do sujeito nulo 168
6.2 O parâmetro do sujeito nulo nas línguas crioulas 170
6.3 O parâmetro do sujeito nulo no português do Brasil 172
6.4 O parâmetro do sujeito nulo no português afro-brasileiro 175
6.4.1 Encaixamento linguístico 177
6.4.2. Encaixamento social 180
6.5 Conclusão 182

7 As sentenças relativas **185**
Ilza Ribeiro

7.1 Tipologia das sentenças relativas 187
7.1.1 Sentenças relativas restritivas e não restritivas 187
7.1.2 Relativas livres e semilivres 188
7.1.3 Relativas existenciais 190
7.1.4 Conclusões parciais 190
7.2 Sentenças relativas e restrições universais 191
7.2.1 Função sintática do pronome relativo na amostra de fala analisada 192
7.2.2 Encaixamento da relativa 193
7.2.3 Estratégias de realização das relativas 193
7.2.3.1 Relativa *pied piping* vs relativa cortadora 193
7.2.3.2 Relativa com pronome lembrete 194
7.2.4 Conclusões parciais 195
7.3 Pronome relativo vs. complementador 197
7.3.1 Distinguindo os relativizadores: pronome relativo vs complementador 197
7.3.2 Tipo de relativizador: especialização de formas 200
7.3.2.1 O pronome QUEM 200
7.3.2.2 A forma relativa O QUE 201
7.3.2.3 O adverbial ONDE 202
7.3.2.4 QUANDO, COMO, CUJO 203
7.3.3 Conclusões parciais 204
7.4 Conclusão 206

8 As construções pseudoclivadas e clivadas **209**
Ilza Ribeiro, Moacir Côrtes Junior

8.1 Tipos de clivagem 210
8.2 Pseudoclivadas e clivadas: estrutura e estatuto informacional 213
8.3 Tipos de pseudoclivadas nas comunidades em estudo 214
8.3.1 Pseudoclivada básica 214
8.3.2 Pseudoclivada invertida 215
8.3.3 Pseudoclivada reduzida 215
8.3.4 Outros casos de focalização do tipo pseudoclivada 216
8.4 Tipos de clivadas nos dados das comunidades em estudo 218
8.4.1 Clivadas básicas 218
8.4.2 Clivadas invertidas 219
8.4.3 Clivadas sem cópula 220
8.4.4 Clivadas com dupla cópula 221
8.4.5 Outros casos de clivagem 222
8.5 Resultados do estatuto discursivo do foco nas comunidades em estudo 223
8.6 Conclusões parciais 225
8.7 Para finalizar, um pouco de estrutura e de história 228

9 As construções de tópico **231**
Edivalda Araújo

9.1 Panorama teórico 231
9.2 Metodologia 233
9.3 Análise dos dados linguísticos 234
9.3.1 Topicalização de Objeto Direto (TOD) 235
9.3.2 Tópico Pendente com Retomada 236
9.3.3 Tópico Cópia 237
9.3.4 Tópico Sujeito 238
9.3.5 Tópico Pendente 240
9.3.6 Tópico com Cópia Pronominal ou Duplo Sujeito 241
9.3.7 Topicalização Selvagem 241
9.3.8 Tópico Locativo 242
9.3.9 Conclusões parciais - dados linguísticos 243
9.4 Análise da relação dos dados linguísticos com os dados sociais 244
9.4.1 Estada fora da comunidade 244
9.4.2 Nível de escolaridade 245
9.4.3 Faixa etária 247
9.4.4 Conclusões parciais - dados sociais 247
9.5 Considerações finais 249

10 A negação sentencial **251**
Rerisson Cavalcante

10.1 Metodologia 251
10.2 Descrição dos dados: aspectos linguísticos 252
10.2.1 Negação sentencial e tipo de frase 253
10.2.2 Negação sentencial e tipo de oração 256
10.2.3 Negação sentencial e realização dos argumentos verbais: complementos verbais 259
10.2.4 Negação sentencial e realização dos argumentos verbais: sujeito 261
10.2.5 Negação sentencial e realização fonética da partícula pré-verbal 263
10.3 Atuação dos aspectos sociais 264
10.4 Conclusões 266

11 A concordância de número **269**
Alan Baxter

11.1 Enquadramento metodológico 270
11.2 Enquadramento teórico 272
11.3 Posição e classe gramatical 274
11.4 Configuração do SN e marcação do plural em dois dialetos afro-lusófonos 276
11.5 O efeito da presença de marcas precedentes sobre a marcação do plural 280
11.6 Variáveis sociais 282
11.7 A relevância dos processos de aquisição de segunda língua e a transferência estrutural 285
11.8 O mecanismo formal de atribuição de plural 289
11.9 Conclusão 293

12 A concordância de gênero **295**
Dante Lucchesi

12.1 O gênero em português 296
12.2 O gênero na história da língua portuguesa 300
12.3 O gênero nos crioulos de base portuguesa 301
12.4 O gênero no português afro-brasileiro 305
12.5 Condicionamentos sócio-históricos da variação na concordância de gênero no português afro-brasileiro 309
12.5.1 A variável faixa etária 309
12.5.2 A variável sexo 313
12.5.3 A variável estada fora da comunidade 314
12.5.4 A variável nível de escolaridade 315
12.5.5 Conclusão da análise do encaixamento social 316
12.6 Conclusão 316

13 O artigo definido **319**
Alan Baxter, Norma Lopes

13.1 Metodologia e hipóteses 320
13.2 Análise dos dados 322
13.2.1 A faixa etária: o perfil diacrônico da variável *artigo definido* 322
13.2.2 Variáveis linguísticas 323
13.2.2.1 Aspectos semânticos do substantivo 323
13.2.2.2 O número do SN 323
13.2.2.3 Presença de outro constituinte capaz de induzir a referência definida 325
13.2.2.4 Interação com o determinante do SN correferencial prévio 327
13.2.2.5 Função sintática do SN 328
13.3 Conclusão 330

14 A concordância verbal **331**
Dante Lucchesi, Alan Baxter, Jorge Augusto Alves da Silva

14.1 A concordância verbal com a primeira pessoa do singular 336
14.1.1 A variável dependente 338
14.1.2 O encaixamento linguístico 339
14.1.3 O encaixamento social 343
14.1.4 Conclusão 345
14.2 A concordância verbal com a terceira pessoa do plural 348
14.2.1 O encaixamento linguístico 349
14.2.2 O encaixamento social 355
14.2.3 Conclusão 359
14.3 A concordância verbal com a primeira pessoa do plural 360
14.3.1 O encaixamento linguístico 364
14.3.1.1 Realização e posição do pronome sujeito 364
14.3.1.2 Paralelismo discursivo 366
14.3.1.3 Saliência fônica 367
14.3.2 O encaixamento social 368
14.3.3 Conclusão 370

15 A concordância em estruturas passivas e de predicativo do sujeito **373**
Dante Lucchesi

15.1 O contato entre línguas e a variação na concordância nominal no português brasileiro 373
15.2 A variação na concordância nominal em estruturas passivas e de predicativo do sujeito no português brasileiro 375
15.3 A variação na concordância nominal em estruturas passivas e de predicativo do sujeito no português afro-brasileiro 376
15.3.1 A variação na concordância nominal de número em estruturas passivas e de predicativo do sujeito no português afro-brasileiro 377
15.3.2 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro 379
15.3.2.1 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro relacionada com a concordância nominal de gênero no SN sujeito 380
15.3.2.2 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro relacionada com a indicação do gênero no SN sujeito 381
15.3.2.3 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro relacionada com a referência ao falante no sujeito da oração 383
15.3.2.4 Encaixamento da variação na concordância de gênero em estruturas passivas e de predicativo do sujeito na estrutura social das comunidades rurais afro-brasileiras 384
15.4 Conclusão 386

16 O emprego do modo subjuntivo **389**
Vivian Meira

16.1 O modo subjuntivo 390
16.2 Análise variacionista do uso do modo subjuntivo no português afro-brasileiro, considerando-se o tipo de oração em que a forma verbal é empregada 391
16.2.1 Análise variacionista do modo subjuntivo nas orações relativas 392
16.2.1.1 Localização temporal do evento expresso na oração relativa em relação ao momento da enunciação 392
16.2.1.2 Tempo do subjuntivo previsto no uso culto 393
16.2.1.3 Morfologia verbal 394
16.2.2 Análise variacionista do modo subjuntivo nas orações completivas 395
16.2.2.1 Tipo da oração em que a completiva está encaixada 396
16.2.2.2 Tempo do subjuntivo previsto no uso culto 397
16.2.2.3 Morfologia verbal 397
16.2.3 Análise variacionista do modo subjuntivo nas orações adverbiais 398
16.2.3.1 Nível de realidade do evento contido na oração adverbial 398
16.2.3.2 Localização temporal do evento expresso na oração adverbial em relação ao momento da enunciação 399
16.2.3.3 Morfologia verbal 400
16.3 Análise de conjunto da variação no emprego do modo subjuntivo no português afro-brasileiro 400

16.3.1 O encaixamento linguístico 401
16.3.2 As variáveis sociais 403
16.3.2.1 Faixa etária 403
16.3.2.2 Nível de escolaridade 404
16.3.2.3 Comunidade de origem 405
16.4 Considerações finais 406

17 O objeto direto anafórico: a categoria vazia e o pronome lexical 409
Cristina Figueiredo

17.1 Explicitando as variantes: uma revisão 411
17.2 Análise linguística 413
17.2.1 O condicionamento semântico 414
17.2.2 O paralelismo discursivo 415
17.2.3 A estrutura da frase 416
17.2.4 Referencialidade do objeto direto anafórico 420
17.2.5 Forma verbal 420
17.2.6 Posição da ocorrência em relação ao antecedente 421
17.3 Condicionamentos sociais 422
17.4 Conclusão 424

18 A alternância dativa 427
Dante Lucchesi, Camila Mello

18.1 A alternância dativa: configurações e escopo 429
18.2 A estrutura de dativo nas línguas crioulas: universais *versus* substrato 435
18.3 A alternância dativa nas comunidades rurais afro-brasileiras isoladas: uma análise sociolinguística 441
18.3.1 Condicionamento linguístico da alternância dativa 443
18.3.2 O encaixamento social da alternância dativa nas comunidades rurais afro-brasileiras isoladas 448
18.4 Considerações finais 453

19 A representação da primeira pessoa do plural 457
Dante Lucchesi

19.1 A gramaticalização de *a gente* e o quadro atual da variação no Brasil 458
19.2 O encaixamento linguístico da implementação de *a gente* no português afro-brasileiro 460
19.2.1 O nível de referencialidade 460
19.2.2 O paralelismo discursivo 462
19.2.3 O tipo de texto 463
19.3 O encaixamento social da implementação de *a gente* nas comunidades de fala analisadas 464
19.3.1 A faixa etária 464
19.3.2 A escolaridade 466
19.3.3 As comunidades 467
19.4 Conclusão 468

20 A flexão de caso dos pronomes pessoais **471**
Dante Lucchesi, Elisângela dos Passos Mendes

20.1 A flexão de caso dos pronomes pessoais na norma culta brasileira 472
20.2 A flexão de caso dos pronomes pessoais nos crioulos de base lexical portuguesa da África 477
20.3 A flexão de caso dos pronomes pessoais no português afro-brasileiro 480
20.4 Conclusão 484

21 O sistema de expressão de posse **489**
Dante Lucchesi, Silvana Silva de Farias Araújo

21.1 Os pronomes possessivos nas línguas crioulas 490
21.2 A expressão da posse em português 495
21.2.1 O sistema de indicação de posse no português brasileiro culto 496
21.2.2 O sistema de indicação de posse no português afro-brasileiro 499
21.3 Análise sócio-histórica da indicação de posse com referência à 1ª pessoa do plural no português afro-brasileiro 501
21.4 *Nosso* e *da gente* no português afro-brasileiro: uma análise variacionista 504
21.4.1 A comunidade 504
21.4.2 O paralelismo formal 505
21.4.3 O tipo de posse 506
21.4.4 A faixa etária 507
21.4.5 Estada fora da comunidade 508
21.4.6 Nível de referencialidade 508
21.5 Considerações finais 509

Conclusão **513**
Dante Lucchesi

Caracterização linguística do português afro-brasileiro 523
Caracterização sociolinguística do português afro-brasileiro 535
Palavras finais 543

Referências **547**

Apresentação

Este livro se originou em uma disciplina sobre línguas *pidgins* e crioulas ministrada por Alan Baxter (na época, Professor Associado da Universidade La Trobe, na Austrália), como professor convidado do Mestrado de *Linguística Portuguesa Histórica*, coordenado pelo Prof. Dr. Ivo Castro, na Universidade de Lisboa. Nessa disciplina, Baxter apresentou a hipótese das origens crioulas do português popular do Brasil, com base nas recentes formulações de Gregory Guy e John Holm. Um de seus alunos, Dante Lucchesi, contestou tal hipótese, em função de sua formação estruturalista, recebida durante a graduação, na Universidade Federal da Bahia (UFBA). Segundo essa visão, as mudanças que afetaram a língua portuguesa no Brasil já estavam prefiguradas na sua estrutura, e o contato com índios e africanos só teria acelerado essas tendências seculares. Apesar da divergência, Alan Baxter orientou o trabalho final da disciplina de Dante Lucchesi sobre os artigos nos crioulos de Cabo Verde e São Tomé, que seria publicado no *Journal of Pidgin and Creole Languages*, em 1993. Além disso, já em meados do ano de 1992, Alan Baxter convidaria Dante Lucchesi para implementar no Brasil o Projeto *Vestígios de Dialetos Crioulos em Comunidades Rurais Afro-Brasileiras Isoladas*, com financiamento do Australian Reaserch Council. Mesmo não estando convencido da hipótese da origem crioula do português popular do Brasil, Dante Lucchesi aceitou de bom grado o convite para atuar numa pesquisa de campo tão significativa, e, já no final do ano de 1992, seria constituída a amostra de fala vernácula das comunidades rurais afro-brasileiras isoladas de Barra e Bananal, no Município de Rio de Contas, na região da Chapada Diamantina, no interior do Estado da Bahia.

No ano seguinte, Alan Baxter e Dante Lucchesi percorreriam o interior dos Estados da Bahia e do Espírito Santo, em busca de comunidades rurais afro-brasileiras isoladas que exibissem uma variedade de português com características mais notavelmente crioulas, sem sucesso. Foram visitadas aproximadamente dez comunidades nesses dois estados que exibiam um falar bem afastado da norma urbana culta brasileira, mas não uma variedade claramente crioulizada do português. Dante Lucchesi ainda visitou e recolheu amostras de fala em comunidades afro-brasileiras no interior do Piauí, com as mesmas características. Além disso, coordenou a recolha de amostras de fala no interior do Estado da Bahia, com uma equipe da Universidade Estadual de Feira de Santana, dando origem ao Projeto de

Estudo da Língua Falada no Semiárido Baiano, implementado nessa universidade pelas Profas. Norma Lúcia de Almeida e Zenaide Carneiro.

Em janeiro de 1994, diante do fato de não haver encontrado características mais radicais nas comunidades visitadas até então, Alan Baxter e Dante Lucchesi organizaram uma recolha na comunidade de Helvécia, no Extremo Sul do Estado da Bahia, contando com a valiosa colaboração do antropólogo Marcos Luciano Lopes Messeder, que já havia contribuído com a recolha nas comunidades de Rio de Contas, e com a participação da estudante de Letras da UFBA Nara Barreto. A comunidade de Helvécia tinha um significado especial, porque nela a dialetóloga Carlota Ferreira havia identificado, no início da década de 1960, o que chamou de "remanescentes de um falar crioulo". As três décadas que separaram as duas pesquisas de campo foram suficientes para que a variedade crioulizada que Carlota Ferreira observou na fala de um casal de idosos da comunidade tivesse desaparecido sem que tivesse sido documentada de forma satisfatória.[1] Apesar disso, foi possível recolher, em 1994, amostras de fala que continham características que ainda individualizam Helvécia mesmo entre as comunidades rurais afro-brasileiras do interior do país, tais como a variação na concordância verbal junto à primeira pessoa do singular e a variação na concordância de gênero no interior do Sintagma Nominal (e.g., *eu trabalha no roça*).

Entre 1993 e 1995, o Projeto *Vestígios de Dialetos Crioulos em Comunidades Rurais Afro-Brasileiras Isoladas* funcionou no Departamento de Letras Vernáculas da Universidade Federal da Bahia, envolvendo a participação de estudantes de graduação em Letras com bolsa de Iniciação Científica do convênio CNPq-UFBA. Com as leituras, observações, reflexões e análises que realizou nesse período, Dante Lucchesi criou a convicção de que, se o contato entre línguas não fora suficientemente intenso para produzir no Brasil uma variedade crioula da língua portuguesa que fosse representativa e duradoura, fora decisivo para desencadear processos de variação e mudança que definiram as características centrais das variedades populares do português brasileiro, particularmente a ampla variação no uso das regras de concordância nominal e verbal.[2] Por outro lado, começou a sistematizar, com base nos fundamentos teóricos da Sociolinguística Variacionista, uma visão da polarização sociolinguística do Brasil, fundamental para a compreensão

[1] Não dispondo de gravadores, Carlota Ferreira realizou apenas anotações fonéticas de frases e expressões produzidas por seus dois informantes. Durante anos tentou organizar uma nova expedição científica a Helvécia sem sucesso.

[2] Não se deve descartar, porém, a possibilidade de ter havido processos localizados e relativamente efêmeros de crioulização do português na história do Brasil, particularmente no século XVII (ver capítulo 1 deste livro). A comunidade de Helvécia pode ter experimentado um desses processos (ver capítulo 2 deste livro).

da formação histórica da realidade linguística brasileira e, particularmente, de como o contato entre línguas afetou cada uma das variedades do português no Brasil. Com essas concepções, Dante Lucchesi suspendeu as suas pesquisas na Bahia, e iniciou, em março de 1996, o seu Doutorado em *Linguística* na Universidade Federal do Rio de Janeiro, sob a orientação do Prof. Anthony Naro.

Tratava-se de uma união surpreendente, pois Anthony Naro e Marta Scherre são as duas grandes vozes que se opõem à hipótese de que o contato do português com as línguas indígenas e africanas foi um fator decisivo na formação histórica da realidade linguística brasileira. Apesar de serem dois dos principais nomes da Sociolinguística Variacionista no Brasil, devendo-se dizer que Anthony Naro é o responsável pela introdução, na década de 1970, desse programa de pesquisa no Brasil, esses dois eminentes pesquisadores têm resgatado a concepção estruturalista de que as mudanças ocorridas no português brasileiro são essencialmente o resultado de uma deriva secular da língua, acelerada no Brasil por fatores externos, como o contato entre línguas.

Mas o respeito mútuo e a abertura para o debate de ideias como elemento crucial na produção do conhecimento permitiram que Dante Lucchesi obtivesse o seu título de doutor, com distinção, em 2000, mesmo defendendo uma hipótese claramente contrária à visão do seu orientador. Isso não significa que a relação entre ambos não tenha sido tensa e marcada por conflitos, o que fez com que Dante Lucchesi percebesse que uma pesquisa sociolinguística e histórica sobre o papel do contato entre línguas no Brasil tem de enfrentar, não apenas as dificuldades objetivas decorrentes da falta de registros históricos e da complexidade das mediações que enformam a relação entre língua e sociedade, mas sobretudo um ambiente muito pouco propício a esse tipo de investigação, em função da hegemonia secular das concepções imanentistas que encerram a língua em sua lógica interna.

De volta à Universidade Federal da Bahia, Dante Lucchesi retomou o seu trabalho de pesquisa, montando o Projeto *Vertentes do Português Rural do Estado da Bahia*, que prevê a ampliação do universo de observação, com a recolha de amostras de fala do português popular não marcado etnicamente. Por um lado, a comparação dessa variedade rural com a variedade das comunidades rurais afro-brasileiras isoladas (o **português afro-brasileiro**) pode produzir evidências empíricas decisivas sobre o papel do contato entre línguas, na medida em que se confirme a hipótese de que os efeitos dos processos de mudança induzidos pelo contato são mais notáveis no português afro-brasileiro do que no português rural. Por outro lado, o contraste entre a fala popular da zona rural e dos pequenos centros urbanos do interior do Estado pode fornecer a base para a constituição de um panorama sociolinguístico do português popular do interior, capturando o processo de nivelamento linguístico a partir da difusão do padrão urbano das grandes cidades para todas as regiões do país.

Não obstante a ampliação do programa de investigação, a pesquisa específica sobre as comunidades rurais afro-brasileiras isoladas continuou predominando nos primeiros anos do novo milênio. Essa pesquisa ganhou um novo impulso com o ingresso de Dante Lucchesi no Programa de Pós-Graduação em *Letras e Linguística* da UFBA. Ao introduzir nesse Programa a disciplina *A participação do contato entre línguas na formação do português do Brasil*, criou condições bastante favoráveis para formar, entre os mestrandos e doutorandos do Programa, pesquisadores preparados teoricamente para enfrentar o tema. Muitos de seus orientandos já eram professores de universidades públicas do Estado e tiveram uma participação decisiva no desenvolvimento da pesquisa. Um deles, Jorge Augusto Alves da Silva, professor da Universidade Estadual do Sudoeste Baiano, foi o responsável pela constituição da amostra de fala da comunidade de Cinzento, no Município de Planalto, na região do Semiárido baiano, próximo a Vitória da Conquista, em 2003. Da mesma forma, Cristina Figueiredo, hoje professora da União Metropolitana de Educação e Cultura (UNIME) e da Universidade Salvador (UNIFACS), realizou, em 2004, a recolha da amostra de fala da comunidade de Sapé, no Município de Valença, com o apoio de Edivalda Araújo, na época professora da Universidade do Estado da Bahia (UNEB), hoje professora da UFBA. As amostras de fala de Sapé e Cinzento, juntamente com as amostras de Helvécia e Rio de Contas, constituem o *Acervo de Fala Vernácula do Português Afro-Brasileiro do Estado da Bahia*, que forneceu a base empírica para todas as análises contidas neste livro. São 48 entrevistas de tipo sociolinguístico com doze membros de cada uma das comunidades estudadas.

Esse avanço da pesquisa também só foi possível em função dos apoios institucionais obtidos e que criaram as condições materiais necessárias à sua implementação. Em janeiro de 2002, Dante Lucchesi recebeu um auxílio do Programa de Apoio à Instalação de Doutores do Estado da Bahia (PRODOC), através de convênio firmado entre a Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB) e a Universidade Federal da Bahia. Os recursos financeiros recebidos possibilitaram a aquisição dos equipamentos utilizados na digitalização do áudio das amostras de fala constituídas em meio magnético (Rio de Contas, Helvécia e Cinzento) e de gravadores digitais, com os quais foi constituída, já em meio digital, a amostra de fala da comunidade de Sapé. Paralelamente, as bolsas de Iniciação Científica concedidas pela FAPESB e pelo CNPq permitiram a integração no Projeto de muitos estudantes de graduação que deram uma contribuição fundamental ao desenvolvimento da pesquisa, realizando atividades de digitalização e transcrição das entrevistas, bem como auxiliando na constituição das amostras de fala. Além disso, Dante Lucchesi passou a receber, desde março de 2003, uma bolsa de Produtividade em Pesquisa do CNPq. Nos anos seguintes, o Projeto *Vertentes* alcançou condições materiais bastante favoráveis de funcionamento ao obter uma sala própria no Instituto de Letras da UFBA, que foi montada com recursos

provenientes do Edital de Infraestrutura n. 001/2003, lançado por um convênio entre a FAPESB e a UFBA. Esses recursos também possibilitaram a confecção da página do projeto na Internet (<http://www.vertentes.ufba.br/>). Com essas condições, foram feitas as recolhas do português popular do interior do Estado, que fornecem a base empírica para a segunda etapa do programa de pesquisa. O *Acervo de Fala Vernácula do Português Popular do Interior do Estado da Bahia* é composto por dois conjuntos de 24 entrevistas de tipo sociolinguístico realizadas com moradores de baixa ou nenhuma escolaridade de dois municípios baianos: Poções e Santo Antônio de Jesus; sendo doze entrevistas feitas com moradores da sede do município, e doze com moradores da zona rural.

Mais recentemente, O Projeto *Vertentes* recebeu novos recursos do CNPq, por meio do Edital CNPq nº 50/2006, para seleção pública de projetos de pesquisa nas áreas de Ciências Humanas, Sociais e Sociais Aplicadas e do Edital MCT/CNPq 15/2007 - Universal. Esses recursos já se destinam à implementação da terceira fase do Projeto, que focaliza, a partir de 2007, o português popular da cidade de Salvador. Já foram realizadas 90 entrevistas em quatro bairros da capital baiana (Liberdade, Plataforma, Itapuã e Cajazeiras) e em um município de sua região metropolitana (Lauro de Freitas). Em seu *corpus* base, o projeto, que passou a denominar-se Projeto *Vertentes do Português Popular do Estado da Bahia*, contará com 60 entrevistas de tipo sociolinguístico, sendo 12 de cada uma das localidades. Estão em curso os trabalhos de transcrição dessas entrevistas, contando mais uma vez com a imprescindível colaboração dos bolsistas de Iniciação Científica e dos mestrandos e doutorandos da UFBA integrados ao Projeto. Os três *corpora* (do português afro-brasileiro, do português popular do interior do Estado e da capital), que reúnem 186 entrevistas com duração média de 50 minutos, formam a maior amostra de fala popular já constituída no país para a pesquisa sociolinguística. Com essa base empírica, o Projeto *Vertentes* pretende elaborar um panorama sociolinguístico do português popular do Estado da Bahia, focalizando os dois grandes vetores de sua configuração sócio-histórica. De um lado, está o contato entre línguas, de que derivam as variedades bastante divergentes que se observam nas comunidades rurais isoladas do interior do país e se espalham, por um *continuum* de variação, até a periferia e os bairros populares das grandes cidades, em função do êxodo rural ocorrido desde meados do século XX. De outro lado, está o processo de nivelamento linguístico que se dá em função da crescente influência das grandes metrópoles brasileiras sobre todas as regiões do país, tendo como grande veículo de difusão os meios de comunicação de massa. Assim, o português popular do Brasil se caracterizaria pela tensão entre esses dois vetores contrários entre si.

A composição de um panorama sociolinguístico do português popular do Estado da Bahia está se completando com dois desdobramentos significativos do Projeto *Vertentes* em outras universidades públicas do Estado da Bahia. Adotando

o desenho metodológico do Projeto *Vertentes* e contando com a assessoria de seu coordenador e auxiliares de pesquisa, professores da Universidade Estadual de Feira de Santana e da Universidade do Sudoeste do Estado da Bahia iniciaram a constituição de amostras do português popular de duas cidades de médio porte do interior do Estado da Bahia: Feira de Santana e Vitória da Conquista. Com esse trabalho, completa-se o conjunto de amostras de todas as variedades do português popular do Estado, que podem ser distribuídas da seguinte maneira, considerando-se distinções independentes relativas à formação sócio-histórica, composição étnica, situação socioeconômica e distribuição espacial: (i) o português afro-brasileiro, constituído a partir da fala de comunidades rurais afro-brasileiras isoladas do interior do Estado (muitas delas oriundas de antigos quilombos); (ii) o português popular das pequenas cidades do interior em oposição à zona rural que as cerca (amostras dos Municípios de Poções e Santo Antônio de Jesus); (iii) o português popular das cidades de médio porte do interior em oposição à zona rural que as cerca (amostras dos Municípios de Feira de Santana e Vitória da Conquista); (iv) o português popular da grande metrópole do Estado e de sua periferia urbana (amostras dos Municípios de Salvador e Lauro de Freitas).

O livro que aqui se apresenta ao público é o corolário de mais de quinze anos de pesquisa sobre a primeira dessas variedades, o **português afro-brasileiro**, conforme definido em seu capítulo 2. O estudo da fala das comunidades rurais afro-brasileiras isoladas se justifica por sua importância histórica, não obstante o seu reduzido peso demográfico atual. Como argumentado na Introdução deste livro, essas comunidades assumem, no plano linguístico, a condição de verdadeiros sítios arqueológicos, fornecendo um precioso testemunho acerca dos processos históricos que marcaram a formação dos padrões de fala da maioria da população do país e que refletem genuinamente o caráter pluriétnico da sociedade brasileira. O amplo conjunto de análises linguísticas precedido de vários capítulos propedêuticos assume, também, um significado especial, na medida em que resgata, no plano do interesse científico, a voz de um dos segmentos mais marginalizados da sociedade brasileira.

Marcadas historicamente por terem a sua origem associada ao abominável processo da escravidão, essas comunidades ainda sofrem com o abandono por parte do Estado brasileiro e pelas relações de exploração econômica que ainda degradam as relações sociais no Brasil, conquanto sua situação de penúria e abandono tenha sido mitigada pelos programas sociais implementados nos últimos anos pelo Governo brasileiro, entre os quais se destacam a Aposentadoria Rural e o Bolsa Família. A perseguição que a mídia conservadora move contra esses programas nada mais é do que o reflexo do profundo desprezo que as elites econômicas brasileiras nutrem pela situação da grande maioria da população do país, submetendo não só os meios de comunicação de massa, como todo o aparelho

de Estado, incluindo o sistema judiciário, em função da manutenção dos seus interesses e privilégios e da perpetuação de uma absurda concentração de renda, mesmo que ao custo de conviver com a verdadeira guerra civil que se trava atualmente na periferia das grandes cidades e no campo brasileiro. Ignorando intencionalmente as raízes sociais do problema, tratam a violência urbana e do campo como caso de polícia, buscando criminalizar movimentos sociais legítimos, como o Movimento dos Trabalhadores Rurais Sem-Terra, o MST, ao tempo em que seus advogados manipulam o Poder Judiciário para manter na impunidade os assassinos dos líderes populares e dos missionários que abraçaram a causa do povo oprimido e explorado. Por acreditar que a ciência não pode fugir aos seus compromissos sociais, este livro se irmana, não apenas às ações afirmativas, que ganharam destaque nos últimos anos, mas a todos os movimentos que lutam pela justiça social e pela verdadeira liberdade neste país, mesmo que contenha apenas análises de aspectos da morfossintaxe da fala das comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia.

A elaboração de tais análises contou com a valiosa contribuição de Ilza Ribeiro, que também recebe bolsa de Produtividade em Pesquisa do CNPq, desde 1996. Especialista em Gramática Gerativa, ingressou na UFBA no ano de 2002 e, logo em seguida, associou-se ao Projeto *Vertentes* e começou a orientar trabalhos de Iniciação Científica, Mestrado e Doutorado que utilizavam o *corpus* do Projeto como sua base empírica. Essa associação foi extremamente importante, na medida em que se assume que a análise dos processos sociais da língua não pode ignorar as determinações gramaticais derivadas do funcionamento da faculdade da linguagem — ponto de vista sistematizado no capítulo 4 deste livro. Portanto, as análises que compõem este volume reúnem os trabalhos de seus organizadores, bem como de seus orientandos. Os capítulos 8 e 10 resultaram, respectivamente, das dissertações de mestrado de Moacir Côrtes Junior e Rerisson Cavalcante orientadas por Ilza Ribeiro, que também orientou a tese de doutorado de Edivalda Araújo, de onde se extraiu o capítulo 9. O capítulo 13 é de autoria de Alan Baxter e de Norma Lopes, sociolinguista brasileira e professora da Universidade do Estado da Bahia, orientada por Baxter em sua pesquisa de doutoramento. Parte do capítulo 14 é proveniente da dissertação de Jorge Augusto Alves da Silva, orientada por Dante Lucchesi, que também orientou as dissertações de mestrado de Vivian Meira, Cristina Figueiredo (esta co-orientada por Ilza Ribeiro) e Silvana Silva de Farias Araújo, que serviram de base, respectivamente, para os capítulos 16, 17 e 21. Os capítulos 18 e 20 têm por base a análise dos dados feita, respectivamente, pelas bolsistas de Iniciação Científica Camila Mello e Elisângela dos Passos Mendes, com a orientação e o acompanhamento de Dante Lucchesi. Todas as orientações de pós-graduação de Dante Lucchesi e Ilza Ribeiro ocorreram no Programa de Pós-Graduação em Letras e Linguística da UFBA.

Esta apresentação objetivou descrever mais de quinze anos de pesquisa empírica e teórica que resultou na publicação deste livro. Tal pesquisa não seria possível sem os apoios institucionais aqui registrados. Dessa forma, fica aqui expresso o reconhecimento e o agradecimento dos organizadores a esses órgãos de fomento à pesquisa, desde o Australian Reaserch Council, que forneceu o financiamento inicial, até a FAPESB e o CNPq, que têm dado um apoio contínuo e de várias modalidades ao Projeto nos últimos anos.

Fica aqui também o agradecimento ao Instituto de Letras e aos colegas que sempre apoiaram esta pesquisa, bem como à EDUFBA, na pessoa da sua diretora, Flávia M. Garcia Rosa, por todo o empenho na edição deste livro. À Profa. Vera Rollemberg, que abraçou este projeto, o reconhecimento por sua rigorosa e competente revisão final dos textos. E a Helder Reis, do ADN Grupo de Lisboa, que gentilmente recuperou a digitalização da foto do casal de idosos de Cinzento, possibilitando o seu uso na capa do livro.

Por fim, deve-se dizer que esta pesquisa contou com a inestimável colaboração dos seus informantes, membros das comunidades estudadas e, particularmente, com a ajuda daqueles que serviram como elo entre os pesquisadores e os locais, como Carmo, membro da Pastoral da Terra e líder das comunidades de Barra e Bananal, em Rio de Contas; e Seu Quéte, um respeitado e bonachão membro da comunidade de Helvécia, que, do alto dos seus setenta anos, levava os pesquisadores aos informantes de Helvécia. Na pessoa desses indivíduos, este livro é dedicado aos **afro-brasileiros**.

Os Organizadores

Os autores

DANTE LUCCHESI é Mestre em Linguística Portuguesa Histórica pela Universidade de Lisboa (1993) e Doutor em Linguística pela Universidade Federal do Rio de Janeiro (2000). Professor Adjunto de Língua Portuguesa da Universidade Federal da Bahia e Pesquisador 1-D do CNPq, coordena o Projeto *Vertentes do Português Popular do Estado da Bahia*. Dentre as suas publicações, destaca-se o livro *Sistema, mudança e linguagem* (Parábola, 2004).

ALAN BAXTER é Mestre em Linguística Hispânica pela La Trobe University (1978) e Doutor em Linguística pela Australian National University (1985). É Professor Catedrático de Linguística da Universidade de Macau e coordenador da área de investigação linguística. As suas publicações incluem os livros *A grammar of Kristang (Malacca Creole Portuguese)* (Pacific Linguistics, 1988), *Maquista Chapado: vocabulary and expressions in Macao's Portuguese Creole* (em co-autoria com Miguel Senna Fernandes; Instituto Cultural de Macau, 2004) e *A dictionary of Kristang (Malacca Creole Portuguese-English)* (em co-autoria com Patrick de Silva; Pacific Linguistics, 2004).

ILZA RIBEIRO é Mestre em Linguística pela Universidade Federal da Bahia (1988) e Doutor em Ciências pela Universidade Estadual de Campinas (1995). Professor Adjunto de Linguística da Universidade Federal da Bahia e Pesquisador PQ-2 do CNPq, desenvolve pesquisa na área de Sintaxe, com especialização em Sintaxe Diacrônica.

CAMILA MELLO é estudante de Letras da Universidade Federal da Bahia e atua no Projeto *Vertentes* desde agosto de 2007 com bolsa de Iniciação Científica do CNPq.

CRISTINA FIGUEIREDO é Mestre e Doutora em Letras pela Universidade Federal da Bahia (2004 e 2009, respectivamente). Também atuou como Professor Substituto de Língua Portuguesa nessa universidade. Atualmente, é professora de Linguística e Língua Portuguesa da União Metropolitana de Educação e Cultura (UNIME) e da Universidade Salvador (UNIFACS). Desenvolve seu trabalho de investigação como pesquisadora associada do Projeto *Vertentes do Português Popular do Estado da Bahia*.

EDIVALDA ARAÚJO é Mestre em Linguística pela Universidade Federal de Minas Gerais, Doutora em Linguística pela Universidade Federal da Bahia. Professor Adjunto da Universidade Federal da Bahia, desenvolve pesquisas na área de Sintaxe.

ELISÂNGELA DOS PASSOS MENDES é Mestre em Linguística pela Universidade Federal da Bahia (2008). Atuou no Projeto *Vertentes* com bolsa de Iniciação Científica no período de agosto de 2003 a julho de 2005.

JORGE AUGUSTO ALVES DA SILVA é Mestre e Doutor em Letras pela Universidade Federal da Bahia (2003 e 2005, respectivamente). Atualmente, é Professor Adjunto de Linguística Românica e Latim da Universidade Estadual do Sudoeste da Bahia, lecionando também a disciplina Sociolinguística Aplicada no Curso de Especialização Teoria e Ensino. É pesquisador do Projeto *Vertentes* e já orientou diversos trabalhos acadêmicos que envolvem a aplicação de pesquisas de campo em práticas pedagógicas.

MOACIR CÔRTES JUNIOR é Mestre em Linguística pela Universidade Federal da Bahia (2006) e professor da Universidade do Estado da Bahia (UNEB).

NORMA LOPES é Mestre pela Universidade Federal da Bahia (1981) e Doutora em Letras e Linguística por essa mesma universidade (2002), com Bolsa Sanduíche em Macau, na China, onde foi orientada por Alan Baxter na área de sócio-história do português, com financiamento do CNPq. É Professor Colaborador da Pós-Graduação em Letras e Linguística da Universidade Federal da Bahia, Professor Adjunto da Universidade do Estado da Bahia (UNEB) e Professor Permanente do Programa de Pós-Graduação em Estudo de Linguagens (PPGEL) da UNEB, onde atualmente representa a linha II Linguagens, Discurso e Sociedade, orientando e sendo responsável pela área de Sociolinguística. Desenvolve pesquisas na área dos estudos da variação do português contemporâneo e a sua contribuição para o entendimento da formação do português brasileiro.

RERISSON CAVALCANTE é Mestre em Letras pela Universidade Federal da Bahia (2007) e doutorando em Semiótica e Linguística Geral pela Universidade de São Paulo, com projeto financiado pela FAPESP. Foi Professor Substituto de Linguística na Universidade Federal da Bahia, professor de Linguística e Língua Portuguesa da União Metropolitana de Educação e Cultura (UNIME) e Professor Visitante de Língua Portuguesa na Universidade de Leiden, Holanda.

SILVANA SILVA DE FARIAS ARAÚJO é professora de Língua Portuguesa da Universidade Estadual de Feira de Santana, onde coordena a Especialização em Estudos Linguísticos, sendo vice-coordenadora do Colegiado do Curso de Pós-Graduação *Lato Sensu* em Letras. Mestre pela Universidade Federal da Bahia (2005), coordena o projeto de pesquisa *A concordância verbal no português falado na região de*

Feira de Santana-BA: elementos para a caracterização da realidade sociolinguística do português do Brasil.

VIVIAN MEIRA é Mestre em Linguística pela Universidade Federal da Bahia (2006) e doutoranda em Linguística pela Universidade Estadual de Campinas. Professora Assistente de Linguística da Universidade do Estado da Bahia (UNEB), desenvolve pesquisas na área de Sintaxe Gerativa e Sociolinguística.

Introdução

Dante Lucchesi

Assiste-se, atualmente, na sociedade brasileira, a um movimento de tomada de consciência de sua condição pluriétnica. Em 1988, a Constituinte, em um ato extremamente significativo, assegurou, no Artigo 68 do Ato das Disposições Constitucionais Transitórias, o direito à propriedade da terra às comunidades rurais afro-brasileiras remanescentes de antigos quilombos. Desde então, vários órgãos do Governo têm-se mostrado sensíveis a essa problemática. Paralelamente, os movimentos negros e indígenas conquistam, a cada dia, mais espaço no cenário político. Recentemente, como medida de maior impacto, o Governo Federal e algumas universidades públicas instituíram cotas étnicas de acesso ao ensino superior, como instrumento de inclusão social de segmentos historicamente marginalizados. E as polêmicas e as controvérsias que tais iniciativas despertaram revelam a tensão e a complexidade que derivam da diversidade étnica da sociedade brasileira. De qualquer forma, não se pode negar que esse movimento reflete a necessidade de uma reparação histórica em relação aos segmentos de indiodescendentes e de afrodescendentes, que, tendo participado ativamente da construção das riquezas materiais e do patrimônio cultural do país, têm sido, ao longo dos séculos, alijados de seus direitos sociais e excluídos dos espaços institucionais e da cidadania.

No plano linguístico, a contribuição dos segmentos indígenas e africanos para a formação da realidade linguística brasileira tem sido menosprezada, ora por razões ideológicas, determinadas por uma visão de "superioridade cultural" do colonizador europeu, ora por opções teóricas imanentistas, que circunscrevem à lógica interna do sistema linguístico as motivações para as suas mudanças. Os obstáculos ideológicos e teoréticos se somam às dificuldades de realizar pesquisas de campo que possam recolher evidências empíricas consistentes da ocorrência no português brasileiro de processos de variação e mudança efetivamente induzidos pelo contato entre línguas, de modo que subsiste a lacuna acerca do real papel dos segmentos indiodescendentes e afro-brasileiros na história linguística do país.

Em seu passado de colônia agroexportadora, o Brasil guarda muitas semelhanças sócio-históricas com outras regiões do continente americano, como o Caribe e o Sul dos EUA, onde emergiram boa parte das **línguas crioulas**

conhecidas atualmente no mundo. Essas línguas resultaram do contato massivo, radical e abrupto desencadeado pelo sequestro e escravização de cerca de dez milhões de africanos, trazidos para o continente americano pelos colonizadores europeus, entre os séculos XVI e XIX. Só para o Brasil, vieram aproximadamente 40% desses africanos escravizados, produzindo situações de contato linguístico muito semelhantes àquelas em que, no Caribe e no Sul dos EUA, vicejaram as línguas crioulas. Entretanto, não há registros conclusivos da existência de uma comunidade crioulófona no Brasil. A única especificidade objetivamente atestada no comportamento linguístico das comunidades rurais afro-brasileiras é a utilização, em algumas delas, de um código de base lexical africana na comunicação intragrupal, em situações muito particulares, como uma espécie de "língua secreta". Mas os antropólogos que documentaram essas línguas secretas afirmam que, nas situações normais de interação linguística, a comunidade usa uma variedade comum de português popular rural.[1]

Assim, apesar de reunir condições sócio-históricas, em princípio, muito propícias à crioulização da língua do colonizador europeu, não ocorreu no Brasil um processo estável, duradouro e representativo de crioulização da língua portuguesa. Mas, se a crioulização do português no Brasil foi, na melhor das hipóteses, um fenômeno historicamente efêmero e localizado, não se pode pensar seriamente que a língua portuguesa não foi diretamente afetada pelo contato do português com as línguas africanas de uma forma bem ampla e representativa, até porque os afrodescendentes se integraram em todos os segmentos sociais e nos mais diferentes ramos da atividade econômica, em todas as regiões do país; concentrando-se, porém, na base da pirâmide social, em função das adversidades históricas que tiveram de enfrentar.

A dimensão do contato linguístico na proporção das situações desencadeadas pelo tráfico negreiro pode ser medida pelo fato de que, até meados do século XIX, os portugueses e seus descendentes diretos constituíam apenas um terço da população brasileira. Os outros dois terços eram constituídos por africanos e índios e seus descendentes, com larga predominância dos primeiros na maior parte do território brasileiro. A grande maioria desses indivíduos tinha de aprender o português nas condições mais adversas, trabalhando como escravos nas grandes plantações do interior do país, de modo que o conhecimento que adquiriam da língua do colonizador se restringia a um vocabulário reduzido, praticamente desprovido de estrutura gramatical. E as crianças que nasciam nessas condições sub-humanas não tinham, normalmente, acesso à língua nativa dos seus pais, muitas vezes falantes de línguas mutuamente ininteligíveis, tendo aquelas de

---

[1] Essa questão é retomada no capítulo 2 deste livro.

desenvolver a sua linguagem a partir do modelo altamente defectivo de português falado como segunda língua por estes.

Essa é a situação prototípica da crioulização. Contudo, é provável que pelo menos uma boa parte dos filhos dos escravos africanos, os chamados *crioulos*, tenha conseguido inserir-se de forma mais favorável na sociedade do colonizador, tendo mais acesso aos modelos da língua dominante, mitigando, assim, a violência da segregação, que determina a gênese das línguas crioulas. Os modelos disponíveis para a transmissão linguística geracional nesses contextos certamente apresentavam *deficits* em relação às situações "normais", em que uma nova geração desenvolve sua língua materna a partir dos modelos fornecidos pela língua nativa dos seus pais. A diferença reside crucialmente no fato de que, na situação "normal", as crianças dispõem, como modelo, de uma língua plena, dotada de todos os seus mecanismos gramaticais, enquanto, nos casos em que ocorre o que aqui se denominará de **transmissão linguística irregular**, as crianças têm de atender aos requerimentos de marcadores de tempo, modo e aspecto, de regência e ligação, operadores pronominais, etc., inerentes ao desenvolvimento de sua língua materna, a partir de dados linguísticos primários que provêm, no caso da maioria dos adultos que as cercam, de uma segunda língua desprovida da maior parte desses elementos e mecanismos gramaticais.

Segundo as mais profícuas teorias acerca da **mudança linguística** atualmente disponíveis, mesmo que esse processo não aconteça em um nível tal de intensidade que faça com que surja, na língua materna da nova geração, uma gramática inteiramente original, como ocorre na **crioulização**, ele certamente produzirá profundas alterações na gramática das variedades linguísticas que dele derivem, pois as mudanças gramaticais encontram o seu momento crítico exatamente na transmissão geracional. Considerando que as variedades de português faladas por pelo menos dois terços da população brasileira passou em sua história por um processo de transmissão linguística irregular (mais ou menos intenso conforme o caso), seria bastante razoável pensar que esse processo estivesse no centro das pesquisas acerca da história linguística do Brasil. Entretanto, a grande maioria das análises publicadas sobre a **história do português brasileiro** sequer menciona o **contato entre línguas** e, quando o fazem, é para negá-lo, fundamentando-se metafisicamente em uma "deriva secular", que regeria uma estrutura linguística hipostasiada e abstraída das relações históricas que a enformam e das suas condições concretas de uso e transmissão. Em sua versão mais extremada, a posição que tem mobilizado seus mais valorosos esforços para refutar a relevância do contato entre línguas na formação das variedades atuais do português brasileiro chega até a negar a própria realidade dos fatos, questionando diferenças qualitativas notáveis entre os processos de variação e mudança que afetaram a língua portuguesa no Brasil e os que têm se sucedido em Portugal, com base na interpretação equivocada

dos testemunhos históricos do português arcaico e de dados fragmentados da pesquisa dialetal portuguesa.

Se não é tarefa difícil demonstrar a falta de consistência teórica e os equívocos metodológicos de observação empírica dessas posições extremadas e cada vez mais anacrônicas em face dos quarenta anos de desenvolvimento da pesquisa sociolinguística, ergue-se como grande desafio superar quase um século da hegemonia do **formalismo** que se estabeleceu na **Linguística Moderna**, desde 1916, com a publicação do *Curso de linguística geral* de Ferdinand de Saussure, cuja grande insígnia era a visão de **língua** como *uma estrutura encerrada em sua lógica interna*. E coloca-se, sobretudo, como tarefa central da historiografia linguística no Brasil superar as limitações dos registros históricos que, em sua grande maioria, se restringem à língua da elite colonial e do Império, de indivíduos que adquiriam o português como língua materna a partir de modelos de falantes nativos dessa língua. Grande parte das análises diacrônicas da língua portuguesa em tempo real tratam apenas do que aconteceu com um terço da população do Brasil, ignorando o que se passou com os outros dois terços de descendentes de africanos e indígenas. A **polarização sociolinguística** que marca a formação histórica da realidade linguística brasileira, apartando a fala de uma elite que sempre teve os olhos voltados para a Europa, em busca de seus modelos culturais e linguísticos, da fala da grande maioria da população que, no cadinho de sua pluralidade étnica, cultural e linguística, forjou os elementos definidores da originalidade cultural e linguística do Brasil, que tanto assombram e encantam o mundo ocidental, desautoriza todos os estudos que apresentam *uma história única para o português brasileiro*. Assim como "**o português são dois**", a sua história é igualmente bifurcada.

O que animou os mais de quinze anos de pesquisa em localidades remotas do interior do país e os esforços empreendidos na elaboração de análises linguísticas consistentes, cujo grande corolário é este livro que agora leva a público as análises de dezesseis aspectos da morfossintaxe do português afro-brasileiro fundamentadas em uma sólida base empírica, foi resgatar a história linguística da grande maioria da população brasileira, que, apesar de serem os protagonistas da construção das riquezas do país e do seu patrimônio cultural, sempre foram discriminados e excluídos, tendo a sua voz calada ou recalcada pelo preconceito que se renova a cada dia, como um mecanismo poderosíssimo de dominação política e ideológica da maioria da população por parte de uma minoria que envida todos os esforços para legitimar como única voz digna aquela que se submeta aos ditames da língua da ex-metrópole. Mesmo à custa de cultivar um profundo complexo de inferioridade linguística, porquanto os modelos de uso do português europeu são em muitos aspectos estranhos à realidade de uso da língua no Brasil, a submissão ao cânone linguístico coimbrão constituiu-se como alternativa histórica das classes dominantes brasileiras para engendrar um poderoso mecanismo ideológico de dominação e exclusão social, dirigindo toda a

sua energia discricionária do preconceito para as marcas linguísticas mais notáveis do caráter pluriétnico da sociedade brasileira.

Este livro se fundamenta na visão de que, no cenário polarizado da formação histórica da realidade linguística brasileira, o contato entre línguas afetou diretamente a formação dos padrões coletivos de fala da maioria da população do país (o que se denomina aqui **norma popular**) e só indiretamente a fala das classes economicamente privilegias, tradicionalmente chamada de **norma culta**. E os mecanismos gramaticais em que os efeitos do contato linguístico são mais notáveis são exatamente os mecanismos da **concordância nominal e verbal**, de modo que a falta de concordância constitui a **grande fronteira sociolinguística** da sociedade brasileira. É sobre ela que incide, em sua forma mais virulenta, o preconceito linguístico. E é ela que é tomada como o parâmetro definidor das capacidades profissionais, intelectuais e políticas de cada indivíduo. A imprensa conservadora reiteradamente invocou a falta de concordância na fala do presidente Lula da Silva para "demonstrar" a sua incapacidade para exercer a Presidência da República. É assim que se manifesta a intolerância dos segmentos mais reacionários diante do fato de um operário governar o país, mesmo que o governo desse operário não comprometa, nem confronte, no fundamental, os interesses dos grandes grupos econômicos nacionais e estrangeiros. Ficam, então, evidentes as raízes históricas do preconceito linguístico, elemento central da ideologia dominante e um poderoso mecanismo na engrenagem da grande exclusão social, econômica e política que esgarça a sociedade brasileira até o paroxismo da violência urbana desenfreada e da miséria que degrada a periferia das grandes cidades do país.

Ao reunir evidências consistentes sobre as relações históricas entre polarização sociolinguística do Brasil e o caráter pluriétnico de sua população, com base na descrição sistemática da gramática de um de seus segmentos mais marginalizados historicamente, as **comunidades rurais afro-brasileiras isoladas**, este livro se integra nas efetivas ações afirmativas de reparação histórica, contribuindo para resgatar a legitimidade da fala popular brasileira em seu caráter pluriétnico. Ampliar o conhecimento sistemático acerca da realidade linguística de setores tão marginalizados da sociedade contribui, por um lado, ao fornecer subsídios para a elaboração de políticas públicas, para que o ensino de língua portuguesa se ajuste mais a realidade da diversidade da língua, tornando-se mais eficaz e democrático, mas esta pesquisa visa, sobretudo, a desarmar o preconceito linguístico, diminuindo os seus nefastos efeitos discricionários. Esses são objetivos deste livro que constitui a primeira grande descrição de conjunto do que se denomina aqui **português afro-brasileiro**, com o significado que se vai explicitar agora.

O conceito de português afro-brasileiro fundamenta-se, não em parâmetros étnicos, mas em parâmetros sócio-históricos. Não se reconhece no Brasil uma fronteira linguística determinada por fatores étnicos, como ocorre, por exemplo,

nos EUA, onde o chamado *Black English* constitui uma variedade específica do inglês empregada pelos afro-americanos. Mas o português afro-brasileiro não é o português empregado pelos afro-brasileiros em geral. Estima-se, por exemplo, que aproximadamente 85% da população da cidade do Salvador seja constituída por afrodescendentes. Entretanto, pode-se dizer que muito provavelmente nenhum deles é falante do português afro-brasileiro, no sentido em que esse termo é empregado aqui. Se o afrodescendente tem curso universitário e é filho de pais também de nível superior, ele certamente será um falante da **norma culta brasileira**. Entretanto, infelizmente, a maioria dos afrodescendentes ainda se situa na base da pirâmide social, sendo geralmente falantes da **norma popular urbana**, ou **rurbana**.

O **português afro-brasileiro** designa aqui uma variedade constituída pelos padrões de comportamento linguístico de comunidades rurais compostas em sua maioria por descendentes diretos de escravos africanos que se fixaram em localidades remotas do interior do país, praticando até os dias de hoje a agricultura de subsistência. Muitas dessas comunidades têm a sua origem em antigos quilombos de escravos foragidos e ainda se conservam em um grau relativamente alto de isolamento. Dessa forma, o português afro-brasileiro guardaria uma especificidade no universo mais amplo do **português popular rural brasileiro** (ou, mais precisamente, **norma popular rural do português brasileiro**), não apenas pelas características sócio-históricas próprias às comunidades em que ele é falado, mas, sobretudo, pelas características linguísticas que o distinguiriam das demais variedades do **português popular do Brasil** (ou melhor, da **norma popular brasileira**).

Um levantamento recente estima que existam cerca de 800 comunidades rurais afro-brasileiras isoladas no Brasil e aproximadamente 250 se situam no Estado da Bahia (ANJOS, 2000). São normalmente pequenos agrupamentos que agregam umas poucas centenas de indivíduos. Nesse sentido, não chegam a 1% da população brasileira, constituindo uma reduzida fração dos 20% da população brasileira que ainda vive no campo. Assim sendo, qual seria a importância de conhecer esse microcosmo do universo sociolinguístico brasileiro?

A grande concentração da população brasileira nos centros urbanos é um fenômeno recente e resulta do grande êxodo rural ocorrido desde meados do século XX. Até o início desse século, mais de 80% da população vivia no campo. O deslocamento massivo da população rural para a periferia das grandes cidades transformou em variação diastrática o que antes era variação diatópica. Tradicionalmente, o campo é mais refratário à normatização linguística, um fenômeno em grande medida urbano. No cenário da polarização sociolinguística do Brasil, a chamada **norma culta** teria se formado e se concentrado nos reduzidos núcleos urbanos fixados quase que somente no litoral. Em oposição às cidades, o português popular brasileiro ia se constituindo no interior do país, onde se instalaram os

engenhos, as plantações de fumo, algodão e mandioca, e onde avançavam a atividade da mineração e a pecuária. Nessa interlândia multilíngue, que concentrava a maioria da população do país formada sobretudo por africanos (e indígenas) e seus descendentes mestiços e endógamos, secundados por um reduzido núcleo de capatazes, senhores, trabalhadores livres e colonos pobres de origem europeia, variedades muito alteradas da língua portuguesa iam-se formando em função da aquisição imperfeita do português como segunda língua em situações muito adversas por parte de uma grande população de falantes adultos e da socialização e nativização desse modelo defectivo de segunda língua, através de sucessivas gerações. Assim, a **norma popular brasileira** atualmente, mesmo em sua variante urbana, exibe ainda os reflexos dos processos de variação e mudança induzidos pelo contato entre línguas que marcaram a sua origem histórica no interior do país.

Porém, se os retirantes levaram para as cidades os seus padrões linguísticos e culturais, foram adquirindo forçosamente, em seu processo de integração, os padrões urbanos de maior valor simbólico, dando origem ao indivíduo **rurbano**, marcado por seu caráter híbrido. Se o objetivo da investigação for identificar os efeitos mais notáveis do contato linguístico, o seu universo de observação deve ser então o interior, e não os centros urbanos. Entretanto, mesmo o panorama sociolinguístico da interlândia brasileira tem mudado muito nas últimas décadas, em função da ampliação da malha rodoviária, da massificação do sistema de educação pública e da poderosa ação dos meios de comunicação de massa. Assim, assiste-se no interior do país a um grande processo de **nivelamento linguístico** com a difusão dos padrões urbanos cultos, ou semicultos, de modo que as marcas mais notáveis do contato linguístico do passado tendem a um progressivo desaparecimento.

Nesse contexto, as comunidades rurais afro-brasileiras isoladas constituem um espaço único para a pesquisa em linguística sócio-histórica que visa a rastrear os reflexos do contato entre línguas na estrutura gramatical das variedades atuais do português brasileiro, pois os efeitos dos processos de transmissão linguística irregular sobre a estrutura gramatical da língua no Brasil seriam mais notáveis exatamente nessas comunidades, em função da combinação das condições históricas em que elas se formaram com o isolamento em que se conservaram até recentemente. Ao reunir essas características, essas comunidades de fala ocupariam um papel específico no cenário das variedades rurais do português brasileiro. Essa especificidade seria atestada com a identificação nelas de processos de variação que estariam ausentes em outras comunidades rurais. Por outro lado, considerando que muitos processos de variação e mudança induzidos pelo contato se alastraram para todas as variedades do português no Brasil, inclusive a sua norma culta, é de se esperar que esses processos se apresentem com um maior nível de intensidade nessas comunidades rurais afro-brasileiras isoladas, mesmo em face das demais comunidades rurais.

Não se pode pensar em qualquer variedade linguística historicamente determinada, sem levar em conta a sua heterogeneidade inerente. Não obstante a sua especificidade, as comunidades rurais afro-brasileiras isoladas não apresentam um comportamento linguístico uniforme. É natural que umas comunidades tenham passado por um processo de transmissão linguística irregular mais intenso em sua formação do que outras e/ou tenham se conservado em um grau maior ou menor de isolamento. Portanto, é natural que os efeitos do contato entre línguas sejam mais notáveis em umas comunidades do que em outras. O que se espera é que, no geral, os efeitos sejam nelas mais notáveis do que nas demais comunidades rurais, o que justificaria o conceito de **português afro-brasileiro** aqui proposto.

Este livro reúne análises linguísticas de dezesseis tópicos da morfossintaxe da gramática de quatro comunidades rurais afro-brasileiras isoladas, situadas em diferentes regiões do interior do Estado da Bahia, com base em amostras de fala vernácula recolhidas entre 1992 e 2004. O conjunto de análises e reflexões aqui reunidas tem por objetivo dar fundamentação empírica à especificidade do português afro-brasileiro como uma variedade do português que foi particularmente afetada pelo contato entre línguas em sua formação histórica. Com isso, este livro pretende colocar no centro da pauta da pesquisa sobre a história sociolinguística do Brasil o contato entre línguas, estabelecendo uma nítida clivagem entre os processos de formação das variedades populares do português brasileiro, ocorridos em contexto de multilinguismo, e os processos de formação da chamada norma culta brasileira, em que os efeitos de mudanças de ordem discursiva ou pressões estruturais se fazem mais presentes, pois essa variedade do português só teria sido afetada de forma indireta pelo contato entre línguas através do **contato dialetal** com as variedades populares. Para cumprir esses objetivos, este livro, dividido em duas partes, se estrutura da seguinte maneira.

A primeira parte reúne cinco capítulos em que são apresentados os fundamentos teóricos e metodológicos da pesquisa, bem como a visão histórica que a embasa. A segunda parte contém dezesseis capítulos, cada um dedicado à análise de um tópico da morfossintaxe da gramática das comunidades de fala estudadas.

No primeiro capítulo do livro, é traçado um **panorama da história do contato entre línguas no Brasil**, com base na seguinte hipótese: se as condições em que o contato do português com as línguas indígenas e africanas ocorreram no Brasil não deram ensejo a processos de crioulização do português duradouros e representativos, elas foram bastante propícias a que processos de variação e mudança desencadeados pelo contato linguístico afetasse, direta ou indiretamente, todas as variedades históricas do português brasileiro, sem atingir a intensidade necessária para produzir uma variedade linguística qualitativamente distinta das demais, pois tal processo (a crioulização) deve ter ocorrido, mas de forma localizada e efêmera,

não produzindo uma variedade linguística que se conservasse em uso até os dias de hoje.

O segundo capítulo aprofunda a **definição de português afro-brasileiro**, fazendo também uma descrição das quatro comunidades rurais afro-brasileiras isoladas, que constituem o universo de observação desta pesquisa, em termos históricos, socioeconômicos e culturais. A constituição de um acervo com amostras de fala vernácula recolhidas em quatro comunidades de regiões diferenciadas do interior do Estado da Bahia foi feita com vistas a capturar a heterogeneidade do português afro-brasileiro, enquanto variedade linguística historicamente determinada.

O terceiro capítulo é dedicado ao conceito de **transmissão linguística irregular**, que se refere fundamentalmente ao processo de socialização e nativização de um modelo defectivo de segunda língua adquirida por uma população de indivíduos adultos, de forma precária, em situações de contato linguístico abrupto, massivo e radical. A reestruturação gramatical desse modelo defectivo, em situações de segregação com acesso restrito aos modelos da língua-alvo, pode gerar uma variedade linguística qualitativamente distinta da língua-alvo, no que se definiu como **pidginização** e **crioulização**. Mas a transmissão linguística irregular não implica necessariamente pidginização/crioulização, podendo resultar na formação de uma variedade histórica da língua-alvo que se caracteriza por exibir processos de variação e mudança induzidos pelo contato entre línguas. A sistematização teórica acerca de como o contato entre línguas afeta a estrutura gramatical da língua-alvo é crucial para que as análises aqui reunidas possam atingir o seu objetivo. Mas ainda persistem muitas lacunas na teoria sobre o contato linguístico, de modo que nem sempre o analista dispõe de uma teoria que defina como o contato afeta aquele aspecto específico da estrutura linguística.

O quarto capítulo aborda outro desafio enfrentado também nas análises aqui apresentadas: utilizar uma **teoria gramatical** consistente em análises variacionistas de padrões coletivos de comportamento linguístico. E no quinto capítulo, que fecha a primeira parte do livro, são apresentados os **procedimentos metodológicos** empregados na constituição da base empírica das análises.

A segunda parte do livro, por sua vez, está subdividida em quatro partes, cada uma correspondendo a um nível da estrutura linguística em que se situam os tópicos analisados: a Sentença, o Sintagma Nominal, o Sintagma Verbal e os pronomes.

No nível da **Sentença**, são analisados: a realização do sujeito pronominal, no capítulo 6, que apresenta uma análise com base na teoria gerativa do parâmetro do sujeito nulo, conjugada com uma abordagem sociolinguística do encaixamento social do processo de variação; as orações relativas, no capítulo 7, no qual o uso de estruturas de relativização entre os membros das comunidades rurais afro-brasileiras isoladas são analisadas segundo o enfoque gerativista, a partir dos princípios da

Gramática Universal; as orações clivadas e pseudoclivadas, no capítulo 8; as construções de tópico, no capítulo 9, que exibe uma detalhada taxonomia das estruturas de tópico encontradas nas comunidades analisadas em confronto com o que já se observou em outras variedades do português brasileiro; e a negação sentencial, no capítulo 10, que identifica as três variantes da negação sentencial no português, bem como o seu condicionamento linguístico e social, nas comunidades rurais afro-brasileiras isoladas.

No nível do **Sintagma Nominal**, são analisadas: a concordância de número, no capítulo 11, que contém um estudo contrastivo entre a variação observada em uma comunidade rural afro-brasileira isolada e uma comunidade de fala formada por descendentes de trabalhadores braçais africanos que foram levados para as empresas agroexportadoras de café e cacau de São Tomé, entre os finais do século XIX e o início do século XX; a concordância de gênero, no capítulo 12, que traça um paralelo entre a variação na concordância de gênero no interior do SN em uma comunidade rural afro-brasileira isolada e a violenta erosão da morfologia de gênero nas línguas crioulas de base lexical portuguesa da África, em contraste com o processo de expansão da morfologia do gênero que se observa na história da língua portuguesa, e apresenta ainda uma ampla análise dos condicionamentos sociais da variação linguística na comunidade de fala analisada; e os artigos, no capítulo 13, que contém uma análise variacionista do emprego do artigo em SNs de referência definida, revelando um processo de mudança aquisicional dessa partícula gramatical, bem como paralelos interessantes com o que se observa nos crioulos de Cabo Verde e São Tomé, no que concerne à ausência do artigo.

No nível do **Sintagma Verbal**, são analisadas: a concordância verbal, no capítulo 14, com a análise de três variáveis linguísticas — a concordância verbal com a 1ª pessoa do singular e com a 1ª e 3ª pessoas do plural —, reunindo evidências empíricas muito significativas em favor da polarização sociolinguística do Brasil e da relevância do contato entre línguas na formação das variedades populares do português brasileiro; o emprego do subjuntivo, no capítulo 15, que revela um processo recente de substituição de formas do indicativo, que se generalizaram no passado em função do contato entre línguas, por formas do subjuntivo em contextos contrafactuais, hipotéticos e de incerteza; a concordância em estruturas passivas e de predicativo do sujeito, no capítulo 16, com novos dados empíricos sobre a polarização linguística do Brasil e os efeitos do contato entre línguas; o objeto direto anafórico de 3ª pessoa, no capítulo 17, com uma análise que procura identificar o estatuto da categoria vazia nas comunidades de fala analisadas; e a alternância dativa, no capítulo 18, que aborda um dos efeitos mais radicais do contato entre línguas nas comunidades analisadas, as construções de objeto duplo, que são gerais entre as línguas crioulas e não fazem parte do repertório linguístico das línguas românicas.

Na parte dedicada aos **pronomes**, as análises têm por objeto: no capítulo 19, a forma do pronome da 1ª pessoa do plural, afetada por um processo de substituição da forma canônica do pronome pessoal *nós* pela expressão *a gente*, que parece estar mais avançada na norma popular e rural do que na norma urbana culta; e a flexão pronominal de caso e o sistema de indicação de posse, nos capítulos 20 e 21, respectivamente; esses dois capítulos revelam como processos distintos e independentes de variação e mudança afetaram a norma popular e a norma culta, configurando historicamente a polarização sociolinguística do Brasil.

A conclusão do livro contém uma sistematização dos resultados das análises realizadas com o objetivo de traçar o perfil linguístico e sociolinguístico do português afro-brasileiro. No **perfil linguístico**, são identificados os processos de variação que seriam privativos dessa variedade específica do português brasileiro, bem como as características que a individualizam nos casos em que a variação atinge as outras variedades linguísticas brasileiras, tendo como corolário estabelecer relações empiricamente motivadas entre os processos de variação e mudança aqui analisados com o contato entre línguas que marca a formação histórica dessas comunidades de fala. O **perfil sociolinguístico** tem por objetivo identificar os processos de mudança em curso nas comunidades de fala analisadas, definindo suas características, tais como: se a mudança é interna à comunidade ou decorre de influências externas; se a variação está situada acima ou abaixo do nível de consciência; e como a mudança está se difundindo na comunidade. O perfil do português afro-brasileiro que se constitui com este livro está na base da próxima etapa desta pesquisa: uma comparação sistemática dos resultados aqui obtidos com os de análise que já estão em curso focalizando outras variedades do português popular do interior do Estado da Bahia, com o objetivo de precisar as correlações empiricamente motivadas entre o contato linguístico e os processos de variação e mudança observados. Com isso, espera-se lançar novas luzes sobre a história sociolinguística do Brasil.

1ª Parte

O contexto sócio-histórico e os fundamentos teóricos e metodológicos

1 História do contato entre línguas no Brasil

Dante Lucchesi

A compreensão da realidade sociolinguística brasileira atual, bem como de sua formação histórica, constitui um amplo campo de pesquisa ainda a ser devidamente explorado pelos sociolinguistas. As significativas diferenças estruturais entre as variedades da língua portuguesa no Brasil e em Portugal, que levaram os linguistas de orientação gerativista a afirmar a existência de gramáticas distintas de cada lado do Atlântico, têm sido objeto de investigações coletivas sistemáticas como as levadas a cabo no âmbito do Projeto *Análise Contrastiva de Variedades do Português* — VARPORT (BRANDÃO; MOTA, 2003). Em aspectos importantes da morfossintaxe, como o paradigma da flexão verbal, a cliticização, o parâmetro do sujeito nulo e o movimento na formação das orações interrogativas, o português europeu conserva as características assentes entre as línguas românicas, enquanto o português brasileiro exibe mudanças que afetam esses parâmetros em função de valores não marcados e que são típicas dos processos de crioulização (ROBERTS, 1997). Por outro lado, sabe-se que, na virada do século XVIII para o século XIX, o português europeu passou por um profundo processo de mudanças fonológicas, com uma violenta redução das suas vogais átonas, que também teve fortes implicações no plano da morfossintaxe, como a fixação da ênclise como colocação pronominal praticamente categórica.

A partir daí, descortina-se o campo sócio-histórico, no qual emergem, para além do plano da análise formal da estrutura linguística, indagações acerca das condições sociais que desencadearam tais mudanças no Brasil e em Portugal. Pagotto (1998, p.53) faz referência à ascensão da burguesia como o contexto histórico que propiciou a difusão do novo padrão fonológico em Portugal. No Brasil, o contato dos colonizadores portugueses com milhões de aloglotas, falantes de mais de mil línguas indígenas autóctones e de cerca de duzentas línguas que vieram na boca de cerca de quatro milhões de africanos trazidos para o país como escravos, é, sem sombra de dúvida, o principal parâmetro histórico para a contextualização das mudanças linguísticas que afetaram o português brasileiro. E processos como esses não devem ser levados em conta apenas para a compreensão das diferenças entre as variedades linguísticas nacionais. O próprio mapeamento

das variedades linguísticas contemporâneas do português europeu e, sobretudo, do português brasileiro, tanto no plano diatópico quanto no plano diastrático, depende crucialmente de uma apurada compreensão do processo histórico de sua formação.

Quando se pensa a realidade linguística do Brasil em seu conjunto, não se pode deixar de ter em linha de conta, com a proeminência que lhe deve ser conferida, o grande abismo que separa uma minoria, que desfruta de bens e serviços do universo da cidadania, da grande maioria, que pouco ou nenhum acesso tem aos bens de consumo, aos serviços sociais e aos direitos sociais básicos — uma situação que coloca o Brasil no rol das nações que elevaram ao paroxismo o exercício da desigualdade e da exploração social. E se levarmos em conta que esse fosso tem sido contínua e persistentemente escavado desde quando se inicia a colonização do Brasil em 1532, temos motivos suficientes para pensar a realidade linguística brasileira como um sistema polarizado, constituído por dois grandes subsistemas — uma norma dita culta e uma norma popular —, cada um deles com sua lógica própria, com suas respectivas tendências de mudança linguística e seu sistema particular de avaliação subjetiva (LUCCHESI, 1998, 2001a, 2002b, 2006a).

A **norma culta** seria, então, constituída pelos padrões de comportamento linguístico dos cidadãos brasileiros que têm formação escolar, atendimento médico-hospitalar e acesso a todos os espaços da cidadania e é tributária, enquanto norma linguística, dos modelos transmitidos ao longo dos séculos nos meios da elite colonial e do Império; modelos esses decalcados da língua da Metrópole portuguesa. A **norma popular**, por sua vez, se define pelos padrões de comportamento linguístico da grande maioria da população, alijada de seus direitos elementares e mantida na exclusão e na bastardia social. Na medida em que grande parte de seus antepassados eram "peças" (seres humanos reduzidos à condição de coisa, para usufruto de seus senhores), deve-se pensar que esses falares se formaram no grande cadinho que fundiu, na fornalha da escravidão em massa, as etnias autóctones e as etnias africanas na forma do colonizador europeu. Dessarte, se é uma variedade da língua do colonizador a que se impõe na fala dos segmentos sociais aí formados, não se pode deixar de perceber as marcas de sua aquisição precária e de sua nativização mestiça.

Assim sendo, o panorama da história sociolinguística do Brasil que se apresenta neste capítulo focaliza as condições históricas que determinaram essa clivagem na realidade linguística brasileira, com particular destaque para as situações de contato entre línguas que estão na base da formação da norma linguística popular, o padrão linguístico atual da grande maioria da população do país.

1.1 Panorama da história sociolinguística do contato entre línguas no Brasil

Os primeiros povoadores portugueses que, no início do século XVI, vieram para o Brasil entraram em contato com as tribos indígenas que habitavam a costa e, por falarem línguas muito aparentadas do tronco tupi, eram capazes de se comunicarem entre si utilizando uma espécie de *koiné*, decalcada da língua tupinambá. Essa língua franca, que viria a ser denominada *língua geral da costa brasileira* — ou simplesmente *língua geral* — foi o instrumento de comunicação adotado pelos portugueses para integrar a força de trabalho indígena inicialmente na extração do pau-brasil e posteriormente no cultivo da cana-de-açúcar, do tabaco e do algodão.

O expediente do escambo para se obter a força de trabalho indígena rapidamente cedeu lugar ao emprego da força, e logo se iniciou o processo de escravização do índio. Paralelamente ao trabalho de submissão material do índio levado a cabo pelos colonizadores portugueses, os missionários da Companhia de Jesus, os jesuítas, realizavam o trabalho de sua submissão espiritual, através da catequese. Para uma maior eficácia na conversão do gentio, os jesuítas também adotaram a língua geral, chegando mesmo a codificá-la e dar-lhe feição escrita, empregando o modelo da gramática portuguesa de então. Desse modo, pode-se perceber que o termo *língua geral* recobre uma diversidade de situações linguísticas:

(i) a *koiné* empregada na comunicação entre as tribos de línguas do tronco tupi da costa brasileira;

(ii) a sua versão como língua franca usada no intercurso dos colonizadores portugueses e indígenas;

(iii) a versão nativizada predominante nos núcleos populacionais mestiços que se estabeleceram no período inicial da colonização; e

(iv) a versão "gramaticalizada" pelos jesuítas sob o modelo do português e utilizada largamente na catequese, até de tribos de língua não tupi — chamados por estes de *tapuias*, que significa 'bárbaro', em tupi.

A sujeição das tribos indígenas foi perpetrada sobremaneira pelos núcleos povoadores de São Paulo, que penetraram pelos sertões de Minas Gerais e do Centro-Oeste, organizando expedições, denominadas *bandeiras*, que mobilizavam vastos contingentes humanos, o que levou a que fossem definidas como "cidades em movimento". A composição étnica da sociedade bandeirante de São Paulo, nesse período, se definiu a partir do cruzamento entre os colonizadores portugueses, na sua esmagadora maioria homens, com as mulheres indígenas. E foram os filhos desses cruzamentos, os chamados *mamelucos*, aqueles que se tornaram

os mais sagazes caçadores de índios, escravizando aqueles que poderiam ser até seus parentes, numa busca obstinada de se afirmar na sociedade dos brancos e superar a sua condição de mestiço, renegando-a (RIBEIRO, 1997).

Sendo majoritariamente mestiça e tendo a esfera doméstica da criação dos filhos, muitos deles bastardos, entregue às mulheres índias, a sociedade paulista foi logo tomada pelo uso massivo da *língua geral*, em sua versão nativizada conforme referido no item (iii) acima.[1] O largo predomínio da língua geral em São Paulo perdurou até o início do século XVIII, quando o advento do ciclo da mineração alterou esse estado de coisas.[2] Situação semelhante se observou também no Maranhão, para onde, entre 1560 e 1580, afluíram os tupinambás afugentados das demais regiões do litoral brasileiro, espalhando-se desde a Serra de Ibiapaba até as regiões banhadas por afluentes do Rio Amazonas.

Por outro lado, as bandeiras difundiram o uso da língua geral pelo interior do país até entre as tribos de língua não tupi, o que define uma quinta situação recoberta pelo termo *língua geral*:

(v) língua franca de base tupi utilizada como língua segunda por tribos de língua não tupi (podendo também nesses casos ocorrer a sua nativização).

Por essa outra via, a língua geral também chega até a Amazônia, sendo adotada por tribos não tupi, basicamente dos grupos aruaque e macro-jê, e lá se mantém em uso, sob uma forma bastante diferenciada (e denominada *nheengatu*, ou seja, 'língua boa'), até meados do século XX, em função da população dessa região ser amplamente constituída por índios e seus descendentes mestiços.[3]

Entretanto, a resistência cultural intrínseca do índio ao trabalho forçado, sobretudo ao trabalho agrícola (que na sua cultura nômade extrativista ocupava uma posição subalterna, sendo entregue às mulheres e crianças), aliada às campanhas contra a escravidão indígena movidas pelos jesuítas, fez com que se fizesse necessário

---

[1] Para A. Rodrigues (1986), a língua geral paulista era diferente da língua geral da costa, embora fosse baseada em uma língua do mesmo tronco linguístico, o tupi, estreitamente aparentada com a língua tupinambá, que servira de base para a formação da língua geral da costa e que foi codificada pelos missionários.

[2] Esse predomínio da língua geral em São Paulo é atestado pelo padre Antônio Vieira, que, em 1694, reconhecia que "as famílias dos portugueses e índios em São Paulo estão tão ligadas hoje umas com as outras, que as mulheres e os filhos se criam mística e domesticamente, e a língua, que nas ditas famílias se fala, é a dos índios, e a portuguesa a vão os meninos aprender à escola" (apud SILVA NETO, 1951 [1963, p.55]). Mas, em 1751, a sua decadência é atestada quando "o padre Manuel da Fonseca falava do uso do Tupi como coisa remota, pois afirmava que 'naqueles tempos [do padre Belchior que evangelizou de 1644 a 1719] era comum a toda Comarca'" (SILVA NETO, 1951 [1963, p.58]).

[3] O *nheengatu* é ainda hoje falado em diversas localidades da região amazônica, tendo-se tornado uma das línguas oficiais do Município de São Gabriel da Cachoeira, do Estado do Amazonas, ao lado da língua portuguesa e das línguas indígenas *tukano* e *baniwa*, através da Lei Municipal nº. 145, de 11 de dezembro de 2002.

buscar uma outra fonte de mão de obra para atender as crescentes demandas dos emergentes e muito prósperos engenhos de cana de açúcar, que começaram a se instalar no Nordeste brasileiro, já no século XVI, e se converteram muito rapidamente no setor econômico mais dinâmico do empreendimento colonial do Brasil.[4] Assim, o sequestro e transporte de populações africanas — que passou à história com a denominação de *tráfico negreiro* — forneceu a força de trabalho para a implementação da cultura agroexportadora do açúcar, que viveria o seu apogeu durante o século XVII.[5] E, certamente, os fabulosos lucros proporcionados pelo tráfico de escravos africanos para servirem aos senhores de engenhos, ávidos por mão de obra para impulsionar o seu próspero empreendimento, falou mais alto do que qualquer resistência cultural indígena ou campanha humanitária de bem intencionados missionários.

Não se sabe ao certo quando os primeiros escravos africanos foram introduzidos no Brasil. Na ausência de informações precisas, Prado Jr. (1974, p.37) alude a referências de sua vinda já na primeira expedição oficial de povoadores, em 1532. Oficialmente, o tráfico negreiro para o Brasil é autorizado por um alvará de D. João III, datado de 29 de março de 1549, facultando aos donos de engenho do Brasil o resgate de escravos da Costa da Guiné e da Ilha de São Tomé, por sua própria conta, até o limite de cento e vinte "peças" para cada engenho montado (RAIMUNDO, 1933, p.26-27). A partir daí a importação de escravos africanos para o Brasil cresce de forma vertiginosa, principalmente em Pernambuco e na Bahia, onde, já no final do século XVI, os africanos ocupavam majoritariamente a base da sociedade colonial brasileira; situação que iria se acentuar no século XVII.[6] Assim, enquanto em São Paulo, no Maranhão e no interior do país (ou seja, nas zonas periféricas da Colônia) a língua geral predominava em função da submissão e aculturação das populações indígenas, a língua portuguesa avançava a partir da Bahia e de Pernambuco, os centros mais dinâmicos e mais intimamente ligados à economia mercantilista, na qual se integrava o projeto colonial brasileiro.[7]

---

[4] Gabriel Soares de Souza, em 1587, registra que havia 50 engenhos de cana-de-açúcar em Pernambuco e 44 na Bahia, contra apenas 3 em São Vicente e 2 no Rio de Janeiro (apud ELIA, 1979, p.50).

[5] Cf. Mattoso (2003, p.54): "Como os índios se revelaram inadaptáveis às tarefas sedentárias, fez-se indispensável aumentar o volume do tráfico africano".

[6] Para Mendonça (1933, p.32-33), "Pernambuco e Bahia, nos dois primeiros séculos da história colonial, são os grandes centros de condensação africana".

[7] Nas palavras de Silva Neto (1951 [1963, p.60]): "ainda se falava a língua geral no Maranhão quando na Bahia, em Pernambuco e no Rio de Janeiro ela já entrava em franca decadência. Em 1653, o infatigável padre Vieira testemunhava que os jesuítas pregavam aos índios na língua deles 'a qual os moradores pela maior parte entendiam'. Outro missionário ilustre diz-nos 'que a usavam os meninos e mulheres portuguesas ainda no confessionário'". Porém, com a introdução da cultura do algodão e do largo contingente de escravos africanos que esta carreou, a língua geral cedeu terreno ao português, deslocando-se para o Pará, como o atesta esse depoimento de 1819: "Hoje, ninguém sabe o tupinambá no Maranhão; mas no Pará o sabem não só os índios cristianizados, mas também muitas pessoas brancas, como eu observei" (apud SILVA NETO, 1951 [1963, p.60-61]).

Nesses dois centros, e depois no Nordeste como um todo, se desenvolverá a cultura agroexportadora da cana-de-açúcar, do algodão e do tabaco (nessa ordem de importância). A cultura do fumo se desenvolverá predominantemente na Bahia, na região de Cachoeira, e desempenhará um importante papel, nos séculos XVIII e XIX, como elemento de troca no comércio de escravos. A cultura do algodão se difundirá nos territórios dos atuais Estados do Ceará, Paraíba, e no alto sertão da Bahia e de Minas Gerais; mas sobretudo no Maranhão, que, apesar da forte presença indígena, receberá largos contingentes de escravos africanos, em razão da já referida resistência cultural indígena ao trabalho forçado e às campanhas movidas em seu favor pelo padre Antônio Vieira. Contudo, foi certamente a lavoura da cana-de-açúcar aquela que deu o tom dos primeiros séculos do período colonial. A partir da empresa agroexportadora do açúcar se constituirá uma sociedade patriarcal articulada em torno da figura do senhor de engenho, e flagrantemente dividida em dois universos: a *casa-grande* e a *senzala*, assim definidos na consagrada equação de Gilberto Freyre (1936, p.XV):

> A casa grande, completada pela senzala, representa todo um sistema econômico, social e político; de produção (a monocultura latifundiária); de trabalho (a escravidão); de transporte (o carro de boi, o banguê, a rede, o cavalo); de religião (o catolicismo de família, com capelão subordinado ao *pater familias*, culto dos mortos etc.); de vida sexual e de família (o patriarcalismo polígamo); de higiene do corpo e da casa (o "tigre", a touceira de bananeira, o banho de rio, o banho de gamela, o banho de assento, o lava-pés); de política (o compadrismo).

Na base de todo esse universo, estava a mão de obra africana escravizada, que se torna cada vez mais representativa na composição demográfica da sociedade açucareira do Nordeste.[8] É possível que os primeiros escravos africanos tenham tido contato com a língua geral, mas, com a redução da presença indígena na zona açucareira, os escravos, desde cedo, passaram a ter contato com o português.[9] Os escravos se dividiam, assim, entre os *ladinos*, que tinham alguma proficiência em português, e os *boçais*, que eram incapazes de se comunicar nessa língua.[10] E, se a

---

[8] Segundo Hasenbalg (1979), a população do Brasil, em 1700, seria de cerca de trezentas mil pessoas, e os brancos seriam, no máximo, cem mil, enquanto os escravos de origem africana, cento e setenta mil.

[9] Para Silva Neto (1951 [1963, p.39]), "esses negros já falavam, certamente, um dialeto crioulo-português, pois a nossa língua foi *geral* nas costas da África durante os séculos XV, XVI e XVII". Já Castro (1990, p.103) considera que "o número de ladinos embarcados era relativamente menor ao número total dos que não falavam português, e, depois, ao número de ladinos que aprenderam a falar o português no Brasil".

[10] Em 1651, o padre Vieira observava que "os escravos e escravas boçais sabiam as orações na língua portuguesa, nam entendendo mais o que dizem, que os papagaios pardos de Angola, ou verdes do Brasil" (apud SILVA NETO, 1951 [1963, p.80]).

língua de intercurso entre colonizadores e escravos era o português, talvez tenha sido possível que, em alguns agrupamentos mais homogêneos, em termos étnicos, os escravos pudessem lançar mão de uma língua franca africana para a comunicação entre eles.[11] O uso da língua geral tupinambá tornava-se, assim, residual.[12]

Por outro lado, a pujança da economia açucareira dará ensejo também a manifestações de requinte,[13] entre as quais, naturalmente, o cultivo de uma língua fortemente ligada aos padrões da Metrópole, que, sob o modelo dos administradores, altos funcionários e autoridades eclesiásticas (todos reinóis), vicejará entre a elite colonial, sobretudo nos ambientes finos dos dois mais importantes centros urbanos da época: Salvador e Recife.[14] Já na casa-grande, nas grandes propriedades rurais, encontram-se os escravos domésticos, que, se possuíam certamente uma maior proficiência em português, não deixavam de influenciar a língua de seus senhores — principalmente pela ação das amas que participavam diretamente da criação dos filhos do seu senhor —, impregnado-lhes a língua com as marcas de sua aquisição imperfeita.[15]

Assim, a expansão da língua portuguesa no Brasil, até início do século XVIII, ocorria paralelamente à expansão da sociedade açucareira do Nordeste, através da fala dos colonos portugueses que afluíam para esse empreendimento e, sobretudo,

---

[11] Cf. Mussa (1991, p.151-152): "Não é seguro opinar sobre a existência de línguas gerais africanas no século XVI; o que se pode afirmar com certeza é que grande parte dos escravos africanos acabavam por aprender o português. Houve, porém, graus diferentes de aprendizagem conforme a posição do escravo no quadro social. Dessa forma os negros ladinos e os escravos domésticos e urbanos tendiam mais a abandonar a língua materna (quando não serviam esporadicamente de 'línguas' ou intérpretes), o que certamente os levaria a expandir o uso do português, enquanto o grupo de negros novos e escravos rurais, maioria como se disse, deviam ter no português apenas uma língua de intercurso com a sociedade branca, sendo admissível que conservassem as respectivas línguas maternas nos vários e pequenos grupos linguísticos que se formavam nas senzalas, guardando para ela algumas funções fundamentais".

[12] Cf. Castro (1990, p.100-101): "a partir do século XVII as línguas indígenas perderam a sua razão de ser como língua veicular nos estabelecimentos da costa, sem dúvida nenhuma, a ser substituídas pelas línguas africanas nas senzalas".

[13] Cf. Elia (1979, p.50): "Em 1574 dizia o jesuíta Fernão Cardim: 'em Pernambuco se acha mais vaidade que em Lisboa'".

[14] O que é atestado nos *Diálogos das grandezas do Brasil*, escrito em 1618: "Haveis de saber que o Brasil é praça do mundo [...] e juntamente academia pública, onde se aprende com muita facilidade toda a polícia, bom modo de falar, honrados termos de cortesia" (apud SILVA NETO, 1951 [1963, p.76]).

[15] Reflexos dessa influência podem ser atestados nas recomendações que, em 1798, o bispo Azeredo Coutinho fazia às mestras de um recolhimento recifense no sentido de que corrigissem os vícios de linguagem das mocinhas provincianas, "principalmente o de pronunciarem as palavras invertendo a ordem das letras, como *breço*, em lugar de *berço* [...]; ou suprimindo no meio das palavras algumas letras, como *teado* em vez de *telhado* [...]; ou, ainda, engolindo a letra última, principalmente no número plural, e nos nomes acabados em agudo, como *muitas flore* em lugar de *muitas flores*, *Portugá* em vez de *Portugal*" (apud SILVA NETO, 1951 [1963, p.72]).

através das variedades defectivas de português adquiridas pelos escravos africanos e transmitidas para os seus descendentes crioulos.

No decorrer do século XVIII, o avanço da língua portuguesa no território brasileiro será ainda mais acentuado, com a descoberta das jazidas de ouro e pedras preciosas nas Minas Gerais, no final do século XVII.[16] A riqueza do solo mineiro era de tal monta que desencadeou uma verdadeira corrida do ouro, provocando uma febre de riqueza que se espalhou como uma epidemia pelo reino, ao ponto de, "em 1732, uma consulta do Conselho Ultramarino registrava a inquietação do Governo nos seguintes termos: 'Por este modo se despovoará o Reino e, em poucos anos, virá ter o Brasil tantos vassalos brancos como tem o mesmo Reino'" (ELIA, 1979, p.99-100). Estima-se que, no decorrer do século XVIII, pelo menos trezentos mil portugueses tenham vindo para o Brasil em busca do enriquecimento rápido.[17] O tráfico negreiro também se orientou para as demandas cada vez maiores de mão de obra para a mineração. Além disso, os senhores de engenho, já decadentes, vendiam os seus escravos para a região das minas, ou para lá se dirigiam com os seus escravos, pois de todo o país afluíam aventureiros em busca da fortuna fácil. Desloca-se também para essa região um grande contingente de pequenos artesãos, pequenos agricultores, profissionais das diversas áreas e principalmente pecuaristas que vão garantir a infraestrutura básica ao empreendimento minerador. Esse constante e massivo deslocamento de populações pobres, que já havia se verificado com o fenômeno das bandeiras, pode explicar em boa medida a homogeneidade diatópica das variedades populares do português do Brasil.

A grande onda migratória vinda de Portugal com o ciclo do ouro certamente favoreceu a difusão da língua portuguesa no Brasil, aumentando o acesso dos escravos aos modelos da língua-alvo do segmento dominante e penetrando nas regiões do interior de São Paulo, Minas Gerais e Goiás, onde antes predominava a língua geral. Portanto, esses processos sociodemográficos podem explicar bem mais o retrocesso da língua geral no século XVIII, do que a tão propalada reforma pombalina e a expulsão dos jesuítas do Brasil, na década de 1750.[18] O avanço da língua portuguesa no território brasileiro — seja em sua variedade nativa, veiculada pelos colonos brasileiros, seja na variedade defectiva, falada pelos escravos africanos e seus descendentes crioulos — dá-se primacialmente sobre uma base socioeconômica, com a expansão das lavouras de açúcar no século XVII e, sobretudo no

[16] Cf. Elia (1979, p.99): "o ouro foi encontrado quase simultaneamente em regiões diversas da zona que é hoje Minas Gerais, entre os anos 1693 e 1695".

[17] Os números neste caso também não são precisos, mas Celso Furtado estima em trezentos mil os portugueses que teriam vindo para o Brasil entre 1700 e 1800 (apud HASENBALG, 1979).

[18] Para uma visão diferente acerca da importância dos atos institucionais do período pombalino para o desenvolvimento da língua no Brasil, veja-se J. H. Rodrigues (1983).

século XVIII, com o grande ciclo das minas, cujo manancial extraordinário de riqueza teve um impacto sobre toda a economia mundial.

Reflexo direto dessa produção de riqueza é o surgimento de centros urbanos importantes, como Vila Rica e São João Del Rey, na região de Minas Gerais. Mas o centro urbano mais beneficiado com a atividade mineira foi o Rio de Janeiro, que passa a ser o maior porto importador de escravos do país[19] e, a partir de 1763, se torna a capital da Colônia e sua cidade mais importante.[20] Obviamente, esse crescimento da vida urbana na Colônia também favorecerá a preponderância do português em suas variedades mais normatizadas, criando-se as condições para uma produção literária mais consistente e uma maior difusão de padrões institucionalizados da cultura e da língua. Porém, esse desenvolvimento urbano (muito aquém do volume de riqueza gerado com a mineração — em função do apetite voraz da Metrópole, que dará azo às primeiras insurreições da elite colonial, como a Inconfidência Mineira, de 1789) — pouco afetará a situação da agora mais numerosa população de africanos e descendentes que se esfalfavam na busca do ouro e dos preciosos diamantes. Esses continuavam a assimilar precariamente o português nas mais aviltantes situações de incivilidade e, em muitos casos, conservaram as suas línguas nativas para a comunicação entre si.

O início do século XIX é marcado pela fuga da Corte Portuguesa para o Brasil, em 1808, no rastro das Campanhas Napoleônicas na Europa. De imediato, deve-se destacar o impacto demográfico da fixação da Corte no Rio de Janeiro. E esse impacto deve ser amplificado por se tratar, não de um contingente de colonos dispersos, mas de um grupo orgânico e que constituía a elite política e cultural da época. Além disso, a vinda da Corte desencadeia todo um processo de mudanças políticas, econômicas e sociais que desembocará na independência política do Brasil, em 1822, e a sua constituição enquanto um estado autônomo. Certamente, todo esse processo teve uma série de reflexos de incremento da urbanização e fortalecimento da cultura institucional, como bem descrito por Serafim da Silva Neto (1951 [1963, p.68-69]):

> Quinze mil pessoas, o séquito da família real, trouxeram para o Rio de Janeiro os hábitos e os esplendores da vida lisboeta. Tudo se foi re-europeizando: a construção das casas, o modo de vestir e, até, de encarar a vida. Melhora-se a topografia, a higiene, o policiamento e a iluminação das cidades.

---

[19] Cf. Mendonça (1933, p.33): "No século XVIII, o Rio semelha um porto africano. O Valongo, mercado de escravos, tem um jeito de Luanda. É a maior feira de escravos de todo o Brasil, que exporta para São Paulo, Minas, Estado do Rio e Goiás".

[20] Para Serrano (1968, p.267): "O movimento do porto acentua-se e, em breve, torna-se o Rio o centro comercial de todo o país. Dispunha, ademais, de um sistema de viação terrestre, fluvial e marítima que podia servir a todas as capitanias. Em 50 anos, de 1750 a 1800, aumentou a população do Rio de Janeiro de 25.000 a cerca de 100.000 habitantes".

> A abertura dos portos faz com que cheguem ao Brasil as comodidades e o luxo do velho mundo. Instala-se a *Régia Tipografia* e funda-se, com as preciosidades de Barbosa Machado, a Biblioteca Nacional.
>
> Rugendas, que esteve no Brasil em 1835, observou a importância do fato: "Basta comparar o número de habitantes do Rio de Janeiro em 1808, com o de agora, para compreender a influência que teve sobre a cidade a chegada da Corte de Portugal. Em 1808, o Rio tinha, no máximo, cinquenta mil habitantes e o número de brancos era, sem comparação, muito inferior ao de pretos. Hoje, a população é de cento e dez mil almas e a desproporção entre negros e brancos muito menor, pois, desde então, se estabeleceram na cidade mais de 24 mil portugueses, sem contar grande número de estrangeiros, principalmente ingleses e franceses, atraídos pela abertura dos portos do Brasil [...]".
>
> De súbito, o Rio de Janeiro tornava-se a capital do Império. E operou-se, em todo o país, uma rápida e profunda urbanização.

Contudo, há de ter-se cautela diante do otimismo do grande filólogo, pois todo esse processo não alterou, em suas linhas mestras, o sistema socioeconômico instalado no Brasil. A economia continuou sendo fundamentalmente agroexportadora, e a força de trabalho continuou sendo majoritariamente escrava. A grande mudança operada no período foi a substituição da mineração pela lavoura agroexportadora do café, como principal fonte da riqueza nacional. As grandes fazendas cafeeiras, que proliferaram, inicialmente, no Vale do Paraíba e, a partir da segunda metade do século, no Planalto Paulista, tornaram-se os principais centros produtores da riqueza do país. Essas regiões passaram a absorver os contingentes de mão de obra escrava das minas exauridas. Com a atividade econômica orientando-se comercialmente para a cidade do Rio de Janeiro, acentuou-se a preponderância econômica (e *a fortiori* cultural e política) dessa Província sobre as demais regiões do país, já estabelecida no ciclo do ouro. Com o continuado declínio da cultura açucareira, manteve-se também o deslocamento de escravos para o Sul do país, a exemplo do que já ocorrera no ciclo da mineração, dando continuidade ao tráfico interno e ao massivo deslocamento das camadas pobres da sociedade.[21]

Mas, passadas as primeiras décadas dos Oitocentos, aparecem os primeiros sinais das grandes mudanças que estavam por vir. Preocupada em ampliar os

[21] Esse deslocamento, em tamanhas proporções e ao longo de tão extenso período de tempo, como foi dito anteriormente, mas deve ser novamente enfatizado, pode explicar uma relativa homogeneidade diatópica da norma popular brasileira, bem como a difusão da influência do contato entre línguas para as mais diversas regiões do país. Vale lembrar também que tal deslocamento da população pobre e mestiça do Brasil se repetiria ao longo do século XX, com os nordestinos migrando para as grandes cidades do Sudeste do país e para as regiões Norte e Centro-Oeste.

mercados consumidores que a Revolução Industrial exigia, a Inglaterra, após ter obrigado Portugal a cessar as suas atividades com o tráfico de escravos, passará a pressionar diretamente o governo brasileiro, após a independência política, em 1822. A primeira consequência dessa pressão será uma lei de 7 de novembro de 1831, que declarava livres todos os escravos que chegassem do estrangeiro a portos brasileiros e caracterizava os traficantes como infratores, sujeitos a penas corporais e multas. Porém essa lei não surtiu efeito, e a prosperidade das fazendas cafeeiras do Vale do Paraíba manterá o tráfico negreiro, em condições semiclandestinas, por mais de uma década.[22] Desse modo, o tráfico negreiro no Brasil só será efetivamente proibido em 1850.

A abolição da escravidão, consequência natural do fim do tráfico, só ocorrerá, entretanto, em 1888 (sintomaticamente seguida, um ano depois, pela Proclamação da República), quando os abolicionistas já haviam conquistado a opinião pública e promoviam a fuga em massa dos cativos das fazendas, e o Exército se recusou a perseguir os escravos foragidos.[23] Só assim, foi dobrada a tenaz oposição do Partido Conservador, o braço político dos grandes proprietários rurais do Vale do Paraíba, que resistiram com todas as suas forças à extinção da escravidão, apesar de o sistema baseado na força de trabalho escrava acarretar um ônus extraordinário aos produtores, favorecendo mais o comércio negreiro e os credores do que os próprios latifundiários — o que em boa medida já havia concorrido para o declínio dos engenhos de açúcar.

Já as novas fazendas que se abriam no interior de São Paulo, apesar de contar inicialmente com a mão de obra escrava, passaram a liderar o processo de mudança do sistema para o trabalho assalariado, que então já se mostrava mais rentável e viável do que o trabalho escravo.[24] Assim, a queda do Império e a proclamação da República correspondem, no plano socioeconômico, ao declínio da economia escravista do Vale do Paraíba e à ascensão ao poder dos grandes fazendeiros de São Paulo e de Minas Gerais, que conseguiram se modernizar, implementando o sistema de trabalho assalariado; situação essa que marcará o

---

[22] De acordo com Pierson (1945, p.89-90), "em 1831, um tratado com a Grã-Bretanha procurou abolir o comércio de escravos. Mas, durante um quarto de século, este acordo teve tão pouca força, que se avalia em quinhentos mil o número de africanos que posteriormente ainda foram trazidos para o Brasil, muitos deles em navios dos Estados Unidos que operavam com capital inglês". Atribui-se a esse fato o surgimento da expressão "para inglês ver".

[23] Sendo o Brasil a última nação do Ocidente a pôr termo oficialmente ao emprego do trabalho escravo. O advérbio *oficialmente* é nesse caso imprescindível, porque a exploração do trabalho no Brasil tem sido tão violenta e aviltante que, até nos dias de hoje, fiscais do Ministério do Trabalho flagram grandes propriedades rurais que mantêm trabalhadores braçais em situações "análogas à escravidão".

[24] Estima-se que, desde 1880, o número de trabalhadores livres em São Paulo já superava o número de escravos.

período da Velha República. De qualquer forma, nem o Império, em seus estertores, nem a República nascente elaboraram qualquer plano para a absorção do enorme contingente de ex-cativos, já que tanto os velhos como os novos latifundiários não tinham nenhum compromisso com qualquer política social; e o movimento abolicionista urbano, fundado numa ainda incipiente opinião pública, pouca interferência tinha nos centros de decisão do país.

Essa situação dos ex-cativos, após o colapso das empresas escravagistas, terá importantes implicações sociolinguísticas. Alguns receberam doações de terra e se dedicaram a uma cultura de subsistência, ficando à margem do grande processo produtivo. Mais marginalizados foram aqueles que abandonaram as grandes fazendas e se deslocaram ainda mais para o interior, em busca de locais ermos, onde pudessem livremente subsistir; confinando com a situação em que já se encontravam as comunidades quilombolas, de escravos foragidos. Se se dirigiam para as cidades, os ex-escravos eram mantidos nas periferias, em situação de grande miséria e marginalidade. Assim, deslocados do grande processo produtivo e abandonados pelas políticas públicas, essas comunidades de ex-escravos formavam nichos, nos quais se conservariam as mudanças decorrentes do contato entre línguas massivo e abrupto; podendo-se imaginar que, nos casos mais radicais das comunidades mais isoladas, continuariam em uso variedades crioulizadas do português, como os remanescentes do "falar crioulo" descobertos, no início da década de 1960, em Helvécia, no Extremo Sul do Estado da Bahia, por Carlota Ferreira (1984), ou mesmo línguas francas de base lexical africana, das quais podem ser um reflexo as línguas secretas, descobertas na década de 1980 nas comunidades de Cafundó, em São Paulo (FRY; VOGT; GNERRE, 1981; VOGT; FRY, 1982, 1983, 1985), e Tabatinga, em Minas Gerais (QUEIROZ, 1984).

Portanto, desde o início do empreendimento colonial, em meados do século XVI, até o final do século XIX, a organização socioeconômica do Brasil pouco se alterou. Tratava-se de uma sociedade predominantemente rural, com modestos núcleos urbanos, que mantinham uma incipiente vida institucional, à qual só uma pequena elite tinha acesso. Já a grande maioria da sua população — constituída por índios, africanos e mestiços — era mantida sob a mais violenta exploração e totalmente segregada da vida institucional ou de qualquer cidadania. Assim, criam-se as condições objetivas para a **polarização sociolinguística do Brasil**.[25] De um lado, nos restritos círculos da elite dos pequenos centros urbanos, "os grandes" da Colônia e do Império cultivavam a língua e as boas maneiras, sob a inspiração dos modelos importados d'além mar. Do outro lado, nas vastas regiões do interior

---

[25] Nos últimos quinze anos, Lucchesi (1994, 1998, 2001a, 2002b, 2006a) vem sistematizando a visão de uma *polarização sociolinguística do Brasil*, a partir dos fundamentos teóricos da Sociolinguística Variacionista.

do país, a língua portuguesa passava por drásticas alterações, sobretudo em função do processo de **transmissão linguística irregular**, desencadeado nas situações de contato entre línguas abrupto, massivo e radical, compreendendo a aquisição precária do português por parte dos índios e africanos, a sua socialização entre esses segmentos e a sua nativização, a partir desses modelos defectivos, entre os descendentes endógamos e mestiços desses índios aculturados e africanos escravizados.[26] Essa polarização que marca a história sociolinguística do Brasil foi destacada já na década de 1950 pelo grande linguista e filólogo Serafim da Silva Neto:

> Dos princípios da colonização até 1808, e daí por diante com intensidade cada vez maior, se notava a dualidade linguística entre a nata social, viveiro de brancos e mestiços que ascenderam, e a plebe, descendente dos índios, negros e mestiços da colônia (1951 [1963, p.88-89]).

Para além do seu vasto conhecimento sobre a história da língua portuguesa, Silva Neto baseou a sua visão em preciosos depoimentos históricos, que através da sua pesquisa vieram à luz. Tal é o caso dessa descrição da situação da língua na Província do Maranhão, feita em 1819, por frei Francisco dos Prazeres:

> Presentemente a língua corrente no país é a portuguesa; os instruídos a falam muito bem; porém entre os rústicos ainda corre um certo dialeto, que, enquanto a mim, é o resultado da mistura das línguas das diversas nações que tem habitado no Maranhão (apud SILVA NETO, 1951 [1963, p.90]).[27]

Assim como esses, vários outros testemunhos históricos fazem referência à mistura de línguas que marca a formação das variedades populares da língua portuguesa no Brasil, nos quatro primeiros anos de sua história.[28] Se não é possível

---

[26] O conceito de *transmissão linguística irregular*, desenvolvido para dar conta das situações de contato entre línguas é o tema do capítulo 3 deste livro. Lucchesi (2008a, 2008b) detalha a relação histórica entre a transmissão linguística irregular e a polarização sociolinguística do Brasil.

[27] Vale lembrar que o Maranhão abrigou uma grande população de índios tupinambás que para lá afluíram, nos finais do século XVI e início do XVII, e recebeu também largos contingentes de escravos africanos importados para o trabalho nas lavouras de algodão, entre os séculos XVII e XIX.

[28] Já no século XVII, dispõe-se deste texto perspicaz do padre Antônio Viera: "[...] falam [as nações asiáticas] a língua portuguesa, mas cada uma a seu modo, como no Brasil os de Angola, e os da terra [...] A língua portuguesa tem avesso e direito: o direito é como nós a falamos, e o avesso como a falam os naturais [...] Eram línguas partidas, não só porque eram muitas línguas, senão porque eram línguas e meias línguas: meias línguas, porque eram meio portuguesas e meio de todas as outras nações que as pronunciavam e as mastigavam a seu modo". Para o século XIX, tem-se o testemunho dos viajantes estrangeiros. Em 1821, Saint-Hilaire "notara que os negros conservavam qualquer coisa de infantil em seus modos, linguagem e ideias... o que é uma das características dos falares crioulos". Em 1868, "o viajante inglês Burton, meticuloso e honesto observador, diz que os escravos negros de Morro Velho falavam luso-hamítico" (apud SILVA NETO, 1951 [1963, p.109]).

determinar com absoluta segurança a ocorrência de processos crioulizantes nesse período, pode-se pensar que o português popular sofreu profundas alterações ao ser adquirido, inicialmente, pelos poucos índios aculturados em português e, posteriormente, por contingentes cada vez mais expressivos de escravos trazidos da África; de modo que é bem razoável imaginar que o português falado por esses segmentos tenha passado por um intenso processo de transmissão linguística irregular, que pode, em determinadas situações mais propícias, ter atingido um nível de pidginização e crioulização.[29] Essas mudanças não chegariam a afetar diretamente o português culto das camadas mais altas da sociedade brasileira, que mantinham estreitos laços culturais e linguísticos com Portugal.

Toda essa situação só começa a se alterar após a abolição da escravidão e o início da imigração de largos contingentes de europeus e asiáticos para o Brasil. A adoção do trabalho assalariado, que amplia o mercado consumidor brasileiro, e os grandes excedentes gerados pela exportação do café criam as condições para a industrialização, implementada nas primeiras décadas do século XX, e intensificada sobretudo a partir do Estado Novo. A industrialização em larga escala desencadeia, a partir da década de 1940, o fenômeno da urbanização do país, levando à formação, já na década de 1970, de verdadeiras megalópoles, como o Rio de Janeiro e São Paulo, que ocupam uma posição reitora na condução da economia e da cultura do país.[30]

No bojo desse processo, são construídas as rodovias que integram todas as regiões do território nacional, e os meios de comunicação de massa experimentam um crescimento vertiginoso, alcançando e influenciando — inicialmente, através do rádio e, depois, através da televisão — até as comunidades mais isoladas dos mais afastados rincões do território brasileiro. Esse gigantesco processo, que alterou profundamente a fisionomia do país (mesmo sem o livrar de suas deformações estruturais), acabou por atenuar, em sua ação niveladora, a acentuada polarização entre a fala das elites urbanas e a fala das populações rurais pobres; tanto ao eliminar, na fala popular, as marcas mais características do processo de transmissão linguística irregular, ocorrido nos séculos anteriores (ou mesmo provocando o desaparecimento, em algumas localidades, de formas incipientes de línguas

---

[29] Cf. Silva Neto (1951 [1963, p.89]): "O grau desse falar crioulizado varia de lugar para lugar: depende da percentagem de brancos e do *status* cultural. Onde menor for o número de brancos, onde a população consistir, quase exclusivamente, de índios, negros ou mestiços, maior será o grau de linguajar crioulizante".

[30] Para se ter uma ideia das proporções desse processo de urbanização da sociedade brasileira, basta que se diga que, em 1940, quase 70% dos 41.236.315 habitantes do Brasil viviam no campo; já em 1970, cerca de 56% da população vivia nas cidades, contra 44% da população rural; e, em 1996, a população urbana corresponde a quase 80% dos 157 milhões de habitantes do Brasil (IBGE, 1987, 1996, 1997).

crioulas); quanto ao permitir que certas mudanças ocorridas na fala popular penetrassem na fala das camadas médias e altas. Nesse caso, o grande fenômeno da imigração ocorrido entre os finais do século XIX e as primeiras décadas do século XX pode ter desempenhado um papel decisivo.

A transição para o trabalho assalariado foi fortemente impulsionada pela vinda para o Brasil de mais de três milhões de imigrantes europeus e asiáticos, entre o final do século XIX e o início do século XX. Esses indivíduos, em sua maioria, ingressaram na base da pirâmide social brasileira, dirigindo-se para o trabalho braçal no campo. Nessas circunstâncias, o modelo mais acessível de que dispunham para a aquisição do português era o proveniente dos capatazes e dos trabalhadores braçais locais, que, em sua maioria, eram ex-escravos africanos ou seus descendentes endógamos e mestiços; ou seja, o português que esses imigrantes — italianos, japoneses, alemães, etc. — aprenderam, ao chegar ao Brasil, era o português popular, com as profundas mudanças decorrentes do processo de transmissão linguística irregular por que este havia passado. Contudo, em função de seu *background* cultural, esses imigrantes ascenderam rapidamente na estrutura social, levando para o seio da norma culta algumas das estruturas de matiz popular que haviam adquirido em seu contato inicial com o português.[31]

Pode-se perceber, assim, o caminho através do qual certos processos de variação e mudança induzidos pelas situações de contato maciço do português com línguas indígenas e africanas puderam penetrar nas camadas médias e altas, generalizando-se no português brasileiro como um todo. Este pode ser o caso da variação na concordância de número (tanto verbal quanto nominal), que teria se originado no processo de transmissão linguística irregular ocorrido na base da pirâmide das sociedades colonial e do Império, e daí se teria difundido, tornando-se hoje um fenômeno que se verifica em todos os extratos da sociedade brasileira (SCHERRE, 1994, p.38); registrando-se, é bem verdade, a enorme diferença de frequência na variação que ainda separa os segmentos de maior poder aquisitivo e nível mais alto de escolaridade da população pobre, analfabeta ou de pouca instrução.[32]

[31] Pode-se pensar aí numa sorte de seleção operada pela reação dos falantes nativos cultos, que seriam mais refratários às estruturas populares mais estigmatizadas e mais tolerantes com as estruturas que colidissem menos com as estruturas mais normais do seu falar. Contudo, como esses julgamentos não se operam no plano abstrato da estrutura linguística, mas nas relações sociais e ideológicas que efetivamente se travam entre os falantes (LUCCHESI, 2004c, p.49-50), pode-se pensar que os falantes nativos das classes média e alta seriam muito menos refratários aos desvios linguísticos de um emergente empresário de origem europeia do que diante dos padrões de fala dos trabalhadores braçais negros e mulatos.

[32] Essa enorme diferença quantitativa no emprego das regras de concordância nominal e verbal constitui um verdadeiro divisor de águas no cenário da polarização sociolinguística do Brasil (LUCCHESI, 2006a, 2007, 2008b).

Por outro lado, ocorreram, também ao longo do século XX, significativas mudanças no panorama cultural e linguístico das camadas mais baixas da população, em função do vasto processo de urbanização da sociedade brasileira, já referido acima, e da crescente influência da cidade sobre o campo. Essa influência cultural e linguística define uma tendência da mudança do português popular em direção aos modelos da norma urbana culta, que atingem e influenciam as camadas mais baixas da população através da televisão, do rádio, ou pelo contato direto, proporcionado pelas modernas condições de transporte, ou mesmo através do precário sistema de ensino público. Nesse contexto, estão em vias de desaparecimento as marcas mais drásticas das alterações produzidas pelo extenso contato entre línguas ocorrido nos séculos anteriores, mantendo-se os vestígios das características mais expressivas do processo de transmissão linguística irregular apenas naquelas comunidades rurais mais isoladas e, dentre essas, naquelas em que o contato entre línguas foi mais radical, e a reestruturação gramatical, mais profunda.

Caso particularmente interessante é o das comunidades rurais afro-brasileiras isoladas, que provêm de antigos quilombos ou de agrupamentos de ex-escravos, que receberam doações de terra no final do século passado.[33] Nessas comunidades, ainda se observa, na fala dos seus membros mais velhos, alguns traços linguísticos que apontam claramente para mudanças drásticas resultantes de um processo pretérito de transmissão linguística irregular. Esses traços praticamente desapareceram na fala dos mais jovens, que, diferentemente de seus avós, ouvem rádio e assistem à televisão, e tiveram acesso a pelo menos um ano de instrução pública. Esse processo de mudanças descrioulizantes pode ter ocorrido de forma muito acelerada no decorrer do século XX, em função do violento processo de penetração no campo das grandes empresas capitalistas.

O cenário inicial é constituído pelo alto grau de isolamento linguístico e cultural em que viviam as comunidades quilombolas, clandestinas e ilegais até a abolição da escravatura. Tal cenário seria altamente favorável à crioulização do português, ou mesmo à manutenção, em uso corrente, de línguas francas de base africana. Essa situação certamente pouco se alterou até as primeiras décadas do século XX. A partir daí, inicia-se o violento processo de ocupação de vastas zonas rurais através dos empreendimentos capitalistas de mineração, extração madeireira e produção de celulose, entre outros. Esses empreendimentos expulsaram de suas terras essas comunidades, e/ou acabaram por integrar os seus indivíduos no processo extrativo e/ou produtivo, através da exploração de sua força de trabalho, alterando drasticamente o seu modo de vida e desarticulando a sua cultura local.

---

[33] As comunidades rurais afro-brasileiras do interior do país, em geral, e as comunidades rurais afro-brasileiras isoladas analisadas neste livro são objeto do próximo capítulo.

A fala dessas comunidades *a fortiori* seria profundamente afetada, perdendo todas as particularidades crioulizantes, ou mesmo desaparecendo, na medida em que a coletividade que a usava se desarticulasse. Desse modo, a intensa e violenta penetração do capitalismo no campo brasileiro pode ter eliminado as eventuais variedades crioulizadas do português no Brasil que tenham sobrevivido até o início do século XX.

Na seção seguinte, busca-se reunir algumas evidências históricas sobre o grau de crioulização do português, focalizando especificamente as condições e os contextos sociolinguísticos em que se deu a integração dos grandes contingentes de escravos africanos na sociedade brasileira.

1.2 O contato do português com as línguas africanas

Os estudiosos que se dedicaram ao tema do contato entre línguas na história sociolinguística do Brasil são quase unânimes em afirmar que o elemento africano desempenhou um papel bem mais relevante no processo de constituição de nossa realidade linguística atual do que o elemento indígena.[34]

Desde os primeiros contatos, as comunidades autóctones brasileiras sentiram o violento impacto da transplantação da civilização cristã ocidental para o Brasil. Com efeito, ao longo desses cinco séculos, perpetrou-se um verdadeiro genocídio dos povos indígenas. Quando não eram chacinados pelas "expedições punitivas", eram dizimados por moléstias adventícias contra as quais não haviam desenvolvido qualquer imunidade — uma simples virose dizimava tribos inteiras. Expulsos de suas terras, buscavam refúgios nas regiões mais recônditas e mais inacessíveis, onde encontravam abrigo da sanha "civilizatória". Quando escravizados, não se adaptavam aos trabalhos forçados, sobretudo aos trabalhos agrícolas, em razão de suas características culturais intrínsecas. Sua adaptação à vida urbana era ainda mais remota.[35]

---

[34] Já em 1933, Mendonça (p.52) refutava "a proeminência indevida que se conferiu ao índio com o prejuízo do negro na formação da nacionalidade brasileira"; proeminência esta que ele atribui à ideologia romântica do "indianismo de Gonçalves Dias, Alencar...", que elegeu o índio mítico como símbolo da nacionalidade, em detrimento da figura proletarizada do negro: "O negro, que sua no eito e, esfalfado, trabalha sob o chicote, não oferece a mesma poesia do índio aventureiro que erra pelas florestas... Se um alicerça obscuramente a economia nacional com a lavoura da cana-de-açúcar e do café, e a mineração do ouro, o outro sugere motivos sentimentais para o passatempo dos elegantes do Império...".

[35] Não é possível estabelecer com precisão o número de índios que viviam no território brasileiro na época do descobrimento. As estimativas variam muito, indo de um a cinco milhões de indivíduos. De qualquer maneira, levando em conta que os recenseamentos mais recentes da população indígena brasileira apontam para números em torno de trezentos mil, pode-se ter uma ideia das dimensões do genocídio que foi perpetrado ao longo de quinhentos anos de contato.

Por essas razões, desde muito cedo, o empreendimento colonial brasileiro esteou-se na força de trabalho dos escravos trazidos da África. Já no século XVII, os africanos constituíam a força motriz do setor mais dinâmico da economia colonial, a cultura agroexportadora do açúcar, no Nordeste. Para além do trabalho braçal produtivo, os africanos eram empregados também em tarefas domésticas e muitos dos serviços urbanos. Chegaram mesmo a desempenhar as funções de controle da produção — como capatazes — ou mesmo de repressão — eram negros muitos dos capitães do mato que caçavam escravos foragidos. Desse modo, ao longo de mais de três séculos, passando pelo ciclo do ouro e atingindo a sociedade cafeeira, os africanos e seus descendentes crioulos[36] garantiram, de forma decisiva, a atividade produtiva do país, participando de quase todas as esferas da vida social; não alcançando, porém, os postos e benefícios de uma cidadania, tão restrita e discricionária.

Portanto, para poder dimensionar a forma como o contato entre línguas se processou ao longo da história da sociedade brasileira, que níveis atingiu e, sobretudo, quais as suas consequências para a formação do panorama linguístico brasileiro contemporâneo, deve-se direcionar o foco para como os contingentes de escravos africanos se integraram à sociedade brasileira e como assimilaram e concorreram para o desenvolvimento da língua que se tornaria hegemônica no Brasil: o português. Nesse sentido, algumas questões de natureza sócio-histórica e demográfica assumem uma importância crucial. Dentre essas, destacam-se as seguintes:

(i) número de escravos africanos trazidos para o Brasil e impacto, em termos demográficos, desse contingente sobre a composição étnica da sociedade brasileira;

(ii) origem geográfica e caracterização etnolinguística dos africanos e sua distribuição pelo território brasileiro;

(iii) formas de interação desses indivíduos, nos níveis social, cultural e linguístico, com o grupo dominante.

Nas próximas seções, busca-se sistematizar algumas das respostas a essas questões.

[36] O adjetivo *crioulo* significava inicialmente 'nascido no lugar'. Desse modo, distinguia-se o *escravo de nação* (o escravo africano) do *escravo crioulo* (nascido no Brasil). Com o fim do tráfico negreiro e o progressivo desaparecimento dos africanos, criaram-se as condições para o processo de evolução semântica, através do qual o termo *crioulo* passou a designar genericamente no Brasil os indivíduos de cor negra.

1.2.1 A proporção de africanos na formação da sociedade brasileira

As estimativas do número de escravos africanos trazidos para o Brasil durante cerca de três séculos de tráfico negreiro variam enormemente. Na ausência de fontes diretas, só é possível chegar a alguma cifra através de processos quantitativos indiretos.[37] Pandiá Calógeras (1927 [1957]) partiu de dados de recenseamentos da escravidão no início do século XIX, para chegar a uma média anual de cinquenta e cinco mil escravos importados. Desse modo, apresentou o número impressionante de treze milhões e quinhentos mil escravos importados, que a maioria dos estudiosos considera exagerado. Roberto Simonsen (1937), baseando-se na produtividade e na vida média do escravo, chegou a um número bem menor: três milhões e trezentos mil. Já Maurício Goulart (1949 [1975]), que adotou um método direto, com base nas poucas informações remanescentes sobre o tráfico, chegou a um número semelhante, que oscilaria entre três milhões e quinhentos mil e três milhões e seiscentos mil. Essa é a cifra que normalmente é tomada por base nos estudos sobre a escravidão no Brasil.[38]

Assim, os números aqui apresentados já indicam um forte impacto dessa população africana na constituição da sociedade brasileira; com suas inevitáveis consequências no plano linguístico. Porém, mais importante do que determinar as dimensões desse impacto em termos absolutos é determinar o peso relativo do segmento de origem africana na composição étnica da sociedade brasileira.

O recenseamento feito pelo padre Anchieta em 1583, quando a importação de escravos africanos ainda estava em seu início, revela a composição étnica da sociedade colonial brasileira apresentada no Quadro 1.

No cômputo geral, os brancos portugueses perfazem cerca de 40% da população integrada na sociedade colonial, no final do século XVI; sendo os restantes 60% distribuídos entre negros e índios.[39] Os índios são mais representativos nas regiões laterais da Colônia, na época: São Vicente, Rio de Janeiro e Espírito Santo. Nos centros econômicos mais dinâmicos e progressistas — Pernambuco e Bahia —, a presença de portugueses, bem como de seus escravos africanos, é mais forte. Os colonizadores brancos constituíam a metade da população

---

[37] A ausência de fontes diretas se deve sobretudo à Circular nº 29 do Ministério da Fazenda, de 13 de maio de 1891, que determinou a queima de todos os arquivos e documentos históricos relativos à escravidão.

[38] Cf. Mattoso (2003, p.53): "Entre a segunda metade do século XVI e 1850, data que assinala a abolição definitiva do tráfico brasileiro, o número de cativos importados é avaliado entre 3.500.000 e 3.600.000. Essas cifras baseiam-se em dados incompletos, mas têm unanimidade entre os que atualmente se voltam para o problema. O Brasil teria, pois, importado 38% dos escravos trazidos da África para o Novo Mundo".

[39] Obviamente, não se leva em conta aqui os índios não aldeados.

na Bahia, sede do Governo-Geral; e 40%, em Pernambuco. Já os escravos africanos somavam quase 20% da população baiana, e constituíam a metade da população em Pernambuco (o número de escravos nascidos no Brasil, os *crioulos*, seria então irrelevante). Essa proporção iria aumentar muito ainda no século XVI (ou já seria mesmo bem maior), pois há os que estimam uma importação de escravos até 1600, na ordem de cem mil pessoas; situação que iria se acentuar no século XVII, com a vinda de seiscentos mil escravos africanos para o Brasil (TAUNAY, 1941, p.256 apud SILVA NETO, 1951 [1963, p.80]).

BRANCOS	NEGROS	ÍNDIOS	TOTAL	
Bahia	12.000	4.000	8.000	24.000
Pernambuco	8.000	10.000	2.000	20.000
São Vicente	1.500	-	1.000	2.500
Rio de Janeiro	750	100	3.000	3.850
Espírito Santo	750	-	4.500	5.250
Porto Seguro	750	-	-	750
Ilhéus	750	-	-	750
Itamaracá	250	-	-	250
Total	24.750	14.100	18.500	57.350

Quadro 1 - População do Brasil por etnia em 1583

Fonte: Silva Neto, 1951 [1963, p.79].

O advento do ciclo das minas no século XVIII intensificará o tráfico negreiro, aumentando o número de africanos na composição étnica brasileira. Entretanto, o segmento que mais cresceu nesse período foi o de colonizadores portugueses, atraídos para o Brasil pela riqueza do ouro. O Quadro 2 apresenta dados do censo realizado em 1798.

GRUPO ÉTNICO	Nº DE HABITANTES
Brancos	1.010.000
Mestiços livres	406.000
Mestiços escravos	221.000
Negros escravos	1.361.000
Índios	250.000
Total	3.248.000

Quadro 2 - População do Brasil por etnia em 1798

Fonte: Azevedo, 1975, p.14-15.

Assim, os colonizadores brancos constituem quase 30% da população, enquanto os escravos negros seriam quase a metade do total de habitantes da Colônia; entre esses "o número de crioulos superava o número de africanos" (MUSSA, 1991, p.160). O número de mestiços cresce muito, graças, sobretudo, à maior presença dos colonizadores brancos. Verifica-se também, no período, um deslocamento populacional do Nordeste para o Sudeste do país. Essa tendência se manterá no século XIX, em função do ciclo do café. Para esse período, dispõe-se dos censos de 1850 e 1890, cujos resultados são apresentados no Quadro 3.

GRUPO ÉTNICO	1850	1890
Brancos	2.482.000	6.302.198
Mestiços	2.732.000	5.934.291
Negros	2.500.000	2.097.426
Índios	302.000	- - -
Total	8.020.000	14.333.915

Quadro 3 - População do Brasil por etnia em 1850 e 1890

Fonte: Chiavenato, 1980, p.237, com adaptações.

Nesse período, os brancos são o grupo que mais cresce em termos absolutos, com o segundo maior crescimento em termos relativos, atingindo mais de 40% do conjunto da população brasileira. Isso se deve ao crescimento vegetativo desse segmento e ao aumento da imigração portuguesa e europeia. Deve-se considerar também os casos de "branqueamento" de mestiços que ascenderam socialmente, sendo o mais expressivo o do grande escritor Machado de Assis, que é caracterizado como mestiço, em sua certidão de nascimento, e como branco, em seu atestado de óbito. Do número total de mestiços, nos dados de 1890, deve ser tirado algo em torno de no máximo 300.000, relativamente ao número de índios aí incluídos. Mesmo assim, o número de mestiços aumenta em quase dez vezes, em menos de cem anos, o que revela as enormes dimensões do fenômeno da miscigenação no Brasil. O número de negros aumenta de 1.361.000 para 2.500.000, entre 1798 e 1850; e cai para pouco mais de 2.000.000, em 1890. O crescimento se deve às proporções espantosas que o tráfico negreiro assumiu na primeira metade do século XIX, apesar das proibições; e o decréscimo resulta do fim do tráfico em 1850. Nesse período, "o número de crioulos já ultrapassava com certeza o número de africanos entre os negros livres ou escravos" (MUSSA, 1991, p.163).

Os dados percentuais, apresentados por Mussa, referentes à composição étnica da sociedade brasileira no período de 1583 a 1890 (Tabela 1) correspondem, em linhas gerais, aos aqui já referidos.

Tabela 1 - População do Brasil por etnia do século XVI ao XIX

ETNIA	1583-1600	1601-1700	1701-1800	1801-1850	1851-1890
Africanos	20%	30%	20%	12%	2%
Negros brasileiros	-	20%	21%	19%	13%
Mulatos	-	10%	19%	34%	42%
Brancos brasileiros	-	5%	10%	17%	24%
Europeus	30%	25%	22%	14%	17%
Índios integrados	50%	10%	8%	4%	2%

Fonte: Mussa, 1991, p.163.

A proporção do segmento de origem africana é o primeiro fator que explica por que não houve no Brasil a formação e estabilização, em níveis demográficos representativos, de uma língua crioula de base portuguesa. A situação mais favorável à crioulização do português se daria com a aquisição precária da língua pelos escravos africanos e a sua nativização, a partir desse modelo defectivo, entre os descendentes crioulos. Note-se que, considerando apenas a proporção de africanos e negros brasileiros, esses dois segmentos só atingem o seu ápice no final do século XVII, quando constituem a metade da população. Com efeito, esse deve ter sido o período em que os processos crioulizantes podem ter ocorrido em níveis mais intensos e representativos, até porque essa proporção pode ter excedido bastante os 50%, sobretudo em algumas zonas açucareiras do Nordeste. Mas, com o aumento da imigração portuguesa nos séculos XVIII e XIX, a proporção de africanos e negros cede terreno para os segmentos branco e mestiço, chegando a constituir apenas 15% do total, em 1890.

1.2.2 Fatores que determinaram a ausência de um processo representativo de crioulização do português no Brasil

O nível de crioulização é inversamente proporcional à facilidade de acesso aos modelos da língua-alvo. Em termos demográficos, o acesso à língua-alvo é determinado pela proporção da população dominante. Em outras palavras, quanto menor for a proporção do segmento dominante, maiores as chances de ocorrer a crioulização. Ora, a proporção de população branca no Brasil nunca foi inferior a 30%, crescendo significativamente no século XIX, quando chega a quase a metade do total. Esse quadro está bastante distante do que deu ensejo às formas mais típicas de crioulização, ocorridas no Caribe. No Haiti e na Jamaica, a proporção de brancos nunca ultrapassou os 10% durante praticamente todo o período da colonização, e o nível de miscigenação entre brancos e negros foi muito mais baixo do que o observado no Brasil. Portanto, de um modo geral, *o acesso aos*

modelos da língua-alvo no Brasil sempre foi maior do que o que se observa nas situações mais típicas de crioulização.[40]

Outro fator que também inibiu a formação de um crioulo brasileiro tem a ver com as condições de vida a que eram submetidos os escravos. A crioulização depende de forma crucial da nativização da língua-alvo, o que, por sua vez, passa pela socialização do uso dessa língua-alvo entre a população dominada. A socialização da língua dos dominantes entre a população dominada e a sua nativização colocam essa língua precariamente adquirida diante de necessidades expressivas e comunicativas novas e que não se apresentavam na simples interação imperativa com os indivíduos do grupo dominante. As lacunas do antigo código de emergência são então preenchidas através da gramaticalização, decalcadas ou não das estruturas originais das línguas do substrato, de formas lexicais da língua-alvo, retidas na aquisição do código de emergência. Esse processo de reestruturação é fundamental nos casos típicos de crioulização.[41]

A situação de vida da maioria dos escravos brasileiros não lhes permitia o acesso a qualquer forma de vida familiar ou social. Em primeiro lugar, em decorrência da alta mortandade: os escravos eram submetidos no eito a condições de exploração tão desumanas, que seu ciclo de vida no trabalho é estimado entre cinco e sete anos. Nessa terrível condição, pode-se imaginar que o escravo não teria a menor condição ou qualquer motivação para constituir uma família, ou mesmo procriar; como demonstra a taxa de crescimento vegetativo negativa da população escrava, que apresentam os historiadores. Além disso, a proporção entre homens e mulheres, entre os escravos trazidos da África, era de três para uma. Some-se a isso, a elevadíssima taxa de mortalidade infantil, que deveria ficar em torno de 60% (MUSSA, 1991, p.155).

A socialização da língua-alvo entre a população escrava também foi prejudicada pelos sucessivos deslocamentos das lavouras do Nordeste para a região das minas, no século XVIII, e dessas para as lavouras de café, no século XIX. A venda de escravos, em muitos casos, desarticulava os agrupamentos já formados, ou mesmo os embriões de núcleos familiares, o que certamente dificultava a socialização.

A crioulização também enfrentou a concorrência do uso de línguas francas africanas entre a população escrava, sobretudo nos ambientes das senzalas. A crioulização depende do preenchimento de lacunas que se dá quando a língua dominante é também usada na interação entre os escravos. Se, nessa interação, os

---

[40] Essa é uma afirmação de caráter geral. Ela não se aplica a certas situações específicas, como, por exemplo, o caso de certas propriedades do interior, em que o número de escravos poderia em muito exceder a proporção geral; ou ainda, o das comunidades de escravos foragidos, os quilombos.

[41] Esse processo é tratado em detalhe no capítulo 3 deste livro.

escravos lançam mão de uma língua franca baseada na língua nativa do segmento étnico africano mais representativo e/ou de maior prestígio, o processo de crioulização é inibido. O uso de línguas francas africanas no Brasil não está totalmente esclarecido, e o tratamento dessa questão conduz a uma ligeira digressão sobre a composição etnolinguística dos escravos africanos trazidos para o Brasil.

O tráfico de escravos africanos para a colonização do Novo Mundo, que teve como principais protagonistas Inglaterra, França, Holanda e Portugal, era feito a partir do litoral africano, onde se instalavam bases militares que recebiam os contingentes de escravos destinados à exportação.[42] A costa ocidental de onde se processava o tráfico estendia-se desde o "delta do Rio Senegal, na atual fronteira entre Senegal e Mauritânia, até o Cabo Frio, ao norte da atual Namíbia". Na costa oriental, a chamada *contracosta*, "a faixa terrestre vai do sul do atual Moçambique até Mogadíscio na Somália" (MUSSA, 1991, p.137-138). "Parece, no entanto, que Moçambique nunca chegou a alcançar cifras proveitosas para o tráfico, devido ao custo das viagens transatlânticas feitas num percurso mais longo, principalmente no que diz respeito ao nordeste brasileiro" (CASTRO, 1990, p.102). De qualquer forma, partindo da costa ocidental ou da costa oriental, "a área do tráfico não se limitava à faixa litorânea, uma vez que os captores de escravos, por serem africanos ou europeus assimilados, avançavam muitas léguas para o interior" (MUSSA, 1991, p.138).

Do ponto de vista etnolinguístico, a região de onde se originavam os escravos trazidos para o Brasil é dividida entre dois grandes grupos: o grupo linguístico kwa (chamados tradicionalmente de *sudaneses*), situado, grosso modo, ao norte da linha do equador, na região do Oeste-Africano; e o banto, que compreende a extensão de terras ao sul do equador. "Os sudaneses apresentam uma grande fragmentação linguística oposta à unidade substancial das línguas banto" (MENDONÇA, 1933, p.16). Nos contingentes de escravos do grupo banto trazidos para o Brasil, predominam "as línguas étnicas majoritárias: o quimbundo, o quicongo e o umbundo". No grupo sudanês, "os seus principais representantes no Brasil foram os nagôs ou iorubás, e os jejes ou povos de língua ewe" (CASTRO, 1990, p.98-99).

A composição etnolinguística dos escravos trazidos para o Brasil também se alterou ao longo dos séculos. No século XVI, há um predomínio dos escravos trazidos da Costa da Guiné sobre os escravos trazidos da região do Congo e de Angola. No século seguinte, irão predominar os escravos de língua banto,

[42] Cf. Mussa (1991, p.136): "Os europeus não adentraram o interior do continente, à exceção dos portugueses em alguns poucos pontos. O comércio escravista funcionava fundamentalmente através da aquisição de escravos capturados por africanos, atividade via de regra controlada pelos soberanos locais. Os escravos recém-capturados ficavam à espera dos navios negreiros em barracões construídos dentro ou nas imediações dessas bases militares, de onde partiam para a América".

nomeadamente o quimbundo e o quicongo, exportados de Luanda, que "se transformou no mais importante porto para o tráfico com o Brasil em geral" (CASTRO, 1990, p.102). No século XVIII, parece ter havido uma divisão do tráfico em duas correntes principais. A primeira, de tráfico de escravos de línguas banto, ligava a região de Angola a Pernambuco e, principalmente, ao Rio de Janeiro, maior porto importador de escravos no período, que os repassava para as outras regiões, sobretudo para Minas Gerais. A segunda rota se estabeleceu a partir da troca do fumo produzido no Recôncavo Baiano com os escravos embarcados na Costa da Mina, do grupo linguístico kwa. Assim, "a Bahia não só teve mão-de-obra escrava em abundância, como manteve quase que o monopólio do tráfico externo com aquela região africana e do tráfico interno dos denominados escravos *minas* para a região dos garimpos, que parece ter absorvido a maioria deles" (CASTRO, 1990, p.106). Essa situação perduraria até a primeira metade do século XIX, quando o tráfico negreiro foi extinto; só que nesse período os escravos minas importados pela Bahia e, principalmente, os escravos de línguas bantos importados pelo Rio de Janeiro eram vendidos para as grandes fazendas de café do Vale do Rio Paraíba e, em menor número, para as emergentes lavouras cafeeiras do interior paulista.

No panorama geral dos três séculos de tráfico, há um grande predomínio de escravos trazidos da zona linguística banto. Os escravos de língua banto são amplamente majoritários mesmo na Bahia, no século XVII, quando o tráfico negreiro assume grandes proporções, estimando-se a importação de mais de meio milhão de indivíduos nesse período (GOULART, 1949 [1975, p.122]). Essa situação só iria se alterar com o estabelecimento da copiosa rota de tráfico ligando a Costa da Mina à Bahia, que angariou a essa Província largos contingentes de falantes de línguas kwa, sobretudo o iorubá. Assim sendo, a grande maioria dos estudiosos é unânime em dividir a influência africana no Brasil entre uma influência predominantemente banto na área do Rio de Janeiro (e no Sudeste como um todo) e na área de Pernambuco para o norte, e uma influência predominantemente iorubá na Bahia.[43]

Esse predomínio banto, sobretudo nos séculos XVI e XVII, reflete-se na formação de línguas gerais africanas no Brasil, de modo que, "nos dois primeiros séculos, o quicongo e o quimbundo, seguidas pelo umbundo, foram as línguas numericamente predominantes na maioria das senzalas ou as de maior prestígio sociológico" (CASTRO, 1990, p.103). Apesar de os proprietários de escravos brasileiros evitarem, por razões de segurança, a homogeneidade etnolinguística

[43] Mattoso (2003, p.23) explica: "há uma tradição, no entanto difícil de provar, de que, se a Bahia preferiu sempre importar os sudaneses, Pernambuco tinha predileção pelos bantos e o Rio de Janeiro selecionava metade de sudaneses e outra metade de bantos".

na sua escravaria,[44] o predomínio dos escravos falantes de línguas bantos, e a semelhança entre essas línguas, deve ter favorecido o uso corrente, durante todo o período da escravidão, de línguas francas de base ora quimbundo, ora quicongo, consoante a predominância de seus falantes fosse na senzala,[45] fosse nos quilombos, onde se encontravam africanos, crioulos e mestiços das mais variadas procedências.[46] Portanto, o veículo da socialização dos escravos segregados na senzala, ou foragidos nos quilombos, pode ter sido, em muitas localidades, não o português precariamente adquirido para o intercurso com os seus senhores, mas uma língua franca de base banto.

É certo que línguas de outros grupos linguísticos africanos também assumiram o estatuto de língua franca no Brasil. A destinação para a região das minas dos escravos falantes de língua do grupo fongbe importados pela Bahia resultou na utilização de uma língua franca de base fon, “que foi atestada, na primeira metade do século XVIII, na região de Vila Rica” (CASTRO, 1990, p.107). Essa língua veicular fon deve ter convivido com outras línguas francas de base quimbundo que provavelmente eram usadas entre os escravos introduzidos pelo porto do Rio de Janeiro.

Com efeito, o predomínio de escravos falantes de línguas bantos no Sudeste deve ter propiciado o uso corrente de línguas francas de base quimbundo entre os escravos de diversas localidades dessa região. Com o tempo, essas línguas foram caindo em desuso, sendo mantidas apenas em situações especiais e muito restritas, e substituídas por variedades de português reestruturadas pelos afrodescendentes. Uma primeira evidência do uso dessas línguas francas africanas foi a descoberta de Aires da Mata Machado Filho, em 1944, de uma língua veicular de base lexical banto, na localidade de São João da Chapada, no Norte de Minas Gerais. Essas línguas chegaram até os dias atuais, em comunidades rurais negras, que as conservam como línguas secretas, e também como uma forma de afirmação de sua

---

[44] Cf. Castro (1990, p.101): “[Nas senzalas], se misturavam africanos de diferentes procedências étnicas a um contingente de indígenas, a fim de evitar rebeliões que pusessem seriamente em perigo a vida de seus proprietários numericamente inferiorizados e em áreas interioranas, isoladas e de difícil acesso, sem grandes comunicações umas com as outras”. E Mattoso (2003, p.22): “a metrópole portuguesa adotou sempre a política de misturar as diferentes etnias africanas, para impedir a concentração de negros de uma mesma origem em uma só capitania”.

[45] Castro (1990, p.101) afirma que “com o domínio banto durante três séculos consecutivos, os dialetos das senzalas de base banto provavelmente foram os mais numerosos e extensos no Brasil”.

[46] É assim que, em relação ao Quilombo dos Palmares, o maior quilombo estabelecido no Brasil (no século XVII, na região de Alagoas), Silva Neto (1951 [1963, p.85]) afirma que: “Acreditamos [...] que os palmarenses falavam um dialeto africano de tipo banto. A razão é a grande maioria dos quilombolas eram angolenses [sic]. A tal ponto que à comunidade dos Palmares chamavam *Angola janga*, isto é Angola pequena”.

identidade étnica. Tal é o caso da *falange*, descoberta na comunidade de Cafundó, em São Paulo, e da *língua do negro da costa*, em Tabatinga, Minas Gerais. Ambas empregam um léxico de base banto (sobretudo quimbundo) com as estruturas gramaticais do português popular.

Se o predomínio de escravos bantos no Sudeste favoreceu o emprego dessas línguas de base quimbundo na região, o predomínio de escravos falantes de línguas do grupo kwa (majoritariamente iorubás) levou à utilização de uma língua franca iorubá (chamada na Bahia *nagô*), que era de uso corrente na Cidade da Bahia no século XIX,[47] devendo ter se prolongado até o início do século XX. No plano da resistência cultural e religiosa, o iorubá converteu-se na língua ritual, nos candomblés da Bahia.

No polo oposto, outros fatores como a integração do negro na sociedade brasileira e, sobretudo, a miscigenação também atuaram contra a formação de uma língua crioula no Brasil. Se a maioria dos escravos se destinava ao trabalho da lavoura, onde a aquisição do português se dava de forma bastante precária, há de levar-se em conta também os escravos que eram destinados aos serviços urbanos e aos trabalhos domésticos. Esses provavelmente tinham um maior acesso aos modelos da língua-alvo e adquiriam uma variedade de português mais próxima da língua de seus senhores. Para além disso, deve ser considerado que a maior proficiência em português sempre trouxe mais vantagens aos escravos, mesmo entre os que se encontravam nos engenhos:

> Ladinos e crioulos gozavam de grande prestígio sociológico na sua situação de bilíngues, isto é, na condição de poder participar de duas comunidades linguisticamente diferenciadas: a da casa-grande e da senzala. Enquanto na casa-grande eram os preferidos para os trabalhos domésticos, privando da sua intimidade, nas senzalas lhes era confiado o poder de disciplina e do comando que se estendia às plantações e aos engenhos, na qualidade de capitães-do-mato e guardas pessoais de seus proprietários (CASTRO, 1990, p.104).

Por outro lado, o surgimento das minas e as proporções que esse ciclo econômico assumiu no Brasil tiveram significativas consequências sociolinguísticas:

> Do ponto de vista da interação social, a condição de escravo da mineração, como a de escravo doméstico e urbano, não se mostrava a mesma do escravo rural. Enquanto nas plantações o trabalho da massa escrava era anônimo e coletivo, sua situação social parecia mais definida, por serem mais rígidas as relações casa-grande e senzala, na região das minas,

[47] Cf. N. Rodrigues (1933 [2004, p.156-157]): "A língua nagô é, de fato, muito falada na Bahia, seja por quase todos os velhos africanos das diferentes nacionalidades, seja por grande número de crioulos e mulatos".

> como nos serviços urbanos e domésticos, o trabalho escravo se destacava por ser fruto de seu esforço individual, o que lhe dava mais possibilidades de comprar ou adquirir a liberdade, tornando-se ele próprio comerciante, traficante, minerador, negro de ofício, a bem dizer, estabelecido num trabalho autônomo (CASTRO, 1990, p.108).

Portanto, a sociedade mineira não só facultava ao escravo uma melhor condição de trabalho, como a possibilidade de ascender socialmente; o que implicava naturalmente uma integração do negro aos padrões culturais (e linguísticos) dominantes. Essa possibilidade de ascensão (e integração) do negro na sociedade se reflete até no imaginário popular com "as lendas sobre escravos que conseguiram angariar fortunas, como a de Chico Rei, de quem se diz *negro angola*, e a de Chica da Silva, *negra mina*" (CASTRO, 1990, p.108). Com efeito, a sociedade que se organiza em torno das minas, com sua mobilidade social decorrente da forma abrupta — ou mesmo contingencial — através da qual a riqueza é produzida (ou melhor, extraída), favorece muito menos a crioulização do que a sociedade agroexportadora, na qual os escravos são segregados como uma massa indistinta, destinada unicamente ao trabalho forçado na lavoura. Uma evidência a favor dessa visão advém do fato de os processos de crioulização mais expressivos da América terem ocorrido nas sociedades agroexportadoras do Caribe; enquanto, no continente, não se verificam processos significativos de crioulização nas regiões que também importaram largos contingentes de escravos africanos para o trabalho nas minas.

Por fim, o fator que foi decisivo para que não tenha ocorrido um processo representativo de crioulização do português no Brasil foi a miscigenação racial. Desde o início da colonização, a miscigenação entre o colonizador europeu e as mulheres índias e negras foi geral e constituiu um dos vetores mais importantes da composição étnica da sociedade brasileira; ao ponto de o contingente de mulatos, no final do século XIX, atingir quase a metade da população do Brasil. O impacto demográfico traz em si a força profunda de um processo sociocultural de enormes proporções. Alargando os níveis de interação sociocultural entre os setores dominantes e dominados, a miscigenação foi descortinando progressivamente, para os indivíduos mestiços, novas vias que lhes permitiam uma maior integração na sociedade.[48] Como consequência disso, temos a posição particular do mestiço,

[48] Cf. Silva Neto (1951 [1963, p.114]): "Em 1686, o rei de Portugal obrigava os jesuítas a reaceitar, em suas escolas, os meninos e os moços pardos, isto é, mulatos. Foi esse um passo decisivo na assimilação do mestiço, pois que as letras os habilitavam a penetrar na tradição do grupo dominante. Em 1774, uma lei conferiu aos pardos acesso 'a todos os ofícios, honras e dignidades, sem discriminação por questão de cor'". E Risério (2004, p.353): "ao longo desse mesmo século XVIII, conquistando sua alforria com maior facilidade, mulatos foram se distanciando mais e mais dos pretos. Foram avançando na hierarquia social".

que, apesar de estar a cavaleiro entre duas culturas, tende historicamente a buscar a sua integração na cultura do grupo dominante.

De fato, a condição social do mestiço e a sua crescente representatividade demográfica lhe proporcionaram condições muito mais favoráveis à integração cultural e à ascensão social, do que as que se ofereciam para o seu ancestral africano ou negro *crioulo*. Se foram poucos os mulatos que, como Machado de Assis e José do Patrocínio, alcançaram a elite da sociedade brasileira já no século passado, pode-se pensar que um expressivo contingente deles estava perfeitamente integrado nas disposições socioculturais hegemônicas.[49] Os reflexos, no plano linguístico, desse esforço do mestiço em se integrar aos padrões culturais da sociedade branca são inegáveis; como atestado na observação perspicaz (posto que um tanto quanto romântica) do conhecido poema de Oswald de Andrade, em que o negro e o bom brasileiro dizem "me dá um cigarro", enquanto o mulato, *mais realista que o rei*, diz "dá-me um cigarro".[50]

Portanto, a miscigenação impediu que, no Brasil, a segregação racial confinasse a população de africanos e seus descendentes em guetos sociais de um mundo cultural à parte, do qual a formação de uma língua crioula e claramente diferenciada (resultante da reestruturação independente da língua do superstrato, com forte influência do substrato) seria a consequência natural. O que ocorreu, e ainda ocorre, no Brasil é um violento processo de segregação social, com evidentes reflexos linguísticos, destacando-se aí as características estigmatizadas das falas populares da população pobre (na sua maioria composta por negros e mestiços), em cuja origem certamente se encontram as mudanças operadas no bojo do processo de transmissão linguística irregular que se inicia com a aquisição defectiva do português por parte de africanos e índios. Entretanto, não ocorreu no Brasil um processo de segregação racial, de que poderia resultar uma variedade linguística autônoma de um segmento afro-brasileiro apartado; nem mesmo se pode falar

---

[49] Cf. Silva Neto (1951 [1963, p.114]): "Ao tempo em que Saint-Hilaire visitou o Brasil, aí por 1820, eram mulatos, na maior parte, os atores de teatro. Cerca de vinte anos depois, informam Kidder e Fletcher, meticulosos observadores da vida brasileira: 'Alguns dos homens mais inteligentes que encontrei no Brasil — homens educados em Paris e Coimbra — eram descendentes de africanos cujos antepassados foram escravos [...] Alguns dos mais assíduos estudantes que aí [na Biblioteca Nacional] se encontram são mulatos'".

[50] Ainda Risério (2004, p.356) afirma: "É mais do que célebre, entre nós, a figura do preto ou do mulato pernóstico. Freyre cita anúncios de escravos fugidos, publicados em nossos jornais oitocentistas, que se referem a casos de escravos mulatos 'muito poetas no falar'. Antes que mero estereótipo racista, o 'mulato pernóstico' é uma entidade que, sobrevivendo ainda hoje, deve ser examinada em pauta sociológica. A empáfia linguística nasce no terreno movediço da busca mulata de símbolos-de-status. 'Se falasse com demasiada simplicidade, talvez as más línguas denunciassem traços da herança materna [negra] em seus versos', disse Roger Bastide, a propósito de Silva Alvarenga. O mulato sempre quis 'falar difícil', porque via a classificação social de quem sabia 'falar difícil'".

hoje de uma variedade do português brasileiro característica de um segmento étnico, um *português negro*, por exemplo (GUY, 1981a, p.324);[51] o que há é um português dito culto, bastante diferenciado do português popular, tanto na frequência de uso de certas variantes linguísticas (nomeadamente as regras de concordância nominal e verbal), quanto no julgamento social do uso de tais variantes, configurando o que se tem denominado a **polarização sociolinguística do Brasil**.

Portanto, os fatores que impediram a ocorrência, na história linguística do Brasil, de um processo de crioulização do português, em níveis socialmente representativos e com uma duração significativa, podem ser sumarizados da seguinte maneira:[52]

(i) a proporção entre a população de origem africana e branca, que proporcionou um maior acesso à língua-alvo do que o observado nas situações típicas de crioulização;

(ii) a ausência de vida social e familiar entre as populações de escravos, provocada pelas condições sub-humanas de sua exploração, pela alta taxa de mortalidade e pelos sucessivos deslocamentos;

(iii) o uso de línguas francas africanas como instrumento de interação dos escravos segregados e foragidos;

(iv) o incentivo à proficiência em português;

(v) a maior integração social dos escravos urbanos, domésticos e das zonas mineradoras;

(vi) a miscigenação racial.

Esses fatores podem explicar por que não se constituiu de forma estável e representativa uma língua crioula no Brasil, apesar da forte presença africana em sua história, mas deve ficar claro que tal afirmação se aplica às situações mais representativas da história sociolinguística do país. Não se pode descartar certas situações laterais específicas, nas quais a forte presença afro-brasileira numa comunidade formada em torno de uma fazenda numa região afastada do interior

[51] Cabe aqui um paralelo com os EUA, onde, por causa da segregação racial, desenvolveram-se variedades reconhecidamente crioulizadas do inglês (como o *gullah*, falado nos Estados da Geórgia e da Carolina do Sul), assim como a variedade popular do inglês falada pelos afro-americanos apresenta marcas sensíveis de um processo anterior de transmissão linguística irregular, apesar da proporção de africanos introduzidos na sociedade americana ter sido bem menor do que a trazida para o Brasil.

[52] Deve-se ter em mente que esses fatores, arrolados aqui em seu conjunto, podem guardar uma relação paradoxal entre si. De um lado, estão os fatores que estão ligados à desagregação das formas de socialização do escravo africano, como é o caso dos fatores (ii) e (iii). De outro lado, estão os fatores decorrentes da socialização dos escravos, sobretudo os crioulos e mulatos, os fatores (iv), (v) e (vi). Portanto, apesar de possuírem efeitos sociais contrários, tais fatores concorrem para um mesmo fim sociolinguístico: inibir processos de crioulização do português.

do país, ou de uma comunidade oriunda de um quilombo, possa ter produzido um intenso processo de transmissão linguística irregular em níveis muito próximos ao da crioulização.

1.3 A transmissão linguística irregular na formação do português brasileiro

Se, nas situações mais representativas da história sociolinguística do Brasil, não ocorreram processos típicos de crioulização, é exatamente aí que se localiza a origem das mais importantes consequências do extenso e maciço contato entre línguas que marca essa história. A aquisição precária do português pelos escravos trazidos da África e pelos índios integrados na sociedade brasileira e a nativização desse modelo defectivo de português como língua segunda nas gerações seguintes de seus descendentes endógamos e mestiços desencadearam um processo de transmissão linguística irregular que teve importantes consequências para a formação da atual realidade linguística brasileira, nomeadamente para as suas variedades populares. Os fatores que inibiram a crioulização passam, assim, a atuar como poderosos fatores de difusão das mudanças induzidas pelo contato entre línguas nas mais diversas comunidades de fala do Brasil. A integração social dos escravos ou dos ex-escravos e, sobretudo, a miscigenação são fatores que atuam em dois sentidos: favorecem a assimilação dos padrões linguísticos dominantes por parte dos dominados, ao tempo em que abrem vias de introdução na fala das camadas médias e altas de estruturas criadas por mudanças ocorridas nos extratos mais baixos.

Portanto, em termos gerais, o processo histórico de constituição da realidade linguística brasileira aponta para a ocorrência de significativas mudanças nas variedades populares do português, em função do contato entre línguas. Contudo, por diversos fatores, relacionados, sobretudo, à maior complexidade da sociedade brasileira *vis-à-vis* às sociedades agroexportadoras do Caribe, por exemplo, essas mudanças não foram de monta a dar ensejo à formação e estabilização de uma língua crioula de base portuguesa, o que pressupõe uma reestruturação original da gramática e/ou a transferência de estruturas das línguas de substrato. Tais processos, se ocorreram, ocuparam uma posição lateral, e suas marcas mais evidentes provavelmente desapareceram no bojo das enormes alterações que se processaram no cenário socioeconômico do país ao longo do século XX.

Desse modo, num nível sociolinguístico mais representativo, deve-se pensar, não em termos de crioulização estrita, mas num processo mais amplo de transmissão linguística irregular, que se caracteriza fundamentalmente pela simplificação e/ou

eliminação de certas estruturas gramaticais; ou ainda, em outras palavras, pelo aumento na frequência de uso das formas não marcadas, bem como a sua generalização paradigmática. Seria esse processo de transmissão linguística irregular de tipo leve que estaria na base da formação das atuais variedades populares do português do Brasil,[53] e a sua marca mais evidente seria a ampla e massiva variação no emprego das regras de concordância nominal e verbal. Porém essa simplificação morfológica, característica das situações de contato entre línguas, não se restringe, no português do Brasil, à eliminação das regras de concordância. Em vários planos da estrutura linguística, observa-se um quadro matizado e complexo, no qual os processos de variação e mudança induzidos pelo contato entre línguas apresentam resultados quantitativos diferenciados nas distintas normas linguísticas que constituem a realidade da língua no Brasil atualmente. Em um inextricável mosaico sociolinguístico, fatores estruturais combinam-se de formas distintas com processos históricos igualmente diferenciados, conforme a variedade linguística do português brasileiro que se focalize.[54]

A pesquisa historiográfica, cujo maior objetivo é deslindar informações seguras que permitam determinar que fatores sociais atuaram na formação de cada variedade linguística brasileira, em meio a um cipoal de lacunas documentais e dados contraditórios entre si, constitui a primeira grande tarefa para aqueles que desejam compreender a realidade linguística brasileira e seus fundamentos históricos. Os determinantes históricos relacionam-se dialeticamente com os achados empíricos das análises variacionistas de aspectos da estrutura linguística das diversas variedades atuais do português brasileiro. Nesse sentido, os resultados da análise histórica constituem importantes balizas para os recortes que se devem proceder sobre a realidade linguística contemporânea, bem como para a interpretação dos dados depreendidos na investigação linguística. A compreensão da polarização da realidade linguística brasileira, bem como de sua formação histórica, desempenha, desse modo, um papel central, pois uma devida compreensão da história sociolinguística do Brasil é fundamental, desde o recorte do objeto de observação empírica até a construção final da análise qualitativa dos processos atualmente em curso na realidade linguística brasileira. Por outro lado, os resultados de análises de

---

[53] Na literatura sobre o tema, também se encontra o termo **semicrioulo** para designar as variedades formadas pelo processo de transmissão linguística irregular de tipo leve (SILVA NETO, 1951 [1963, p.107]). Mais recentemente, Holm (2004) desenvolveu o conceito de **reestruturação parcial**, que é análogo ao de transmissão linguística irregular de tipo leve. Holm reúne como variedades linguísticas que se formaram nesse tipo de contexto: o inglês afro-americano dos Estados Unidos, o espanhol caribenho, o português popular do Brasil, o *Afrikaans* (que se desenvolveu a partir do holandês na África do Sul) e o francês da Ilha de Reunião.

[54] O capítulo 3 deste livro apresenta uma análise detalhada do processo de transmissão linguística irregular, em geral, e da sua ocorrência especificamente no Brasil.

aspectos estruturais de variedades contemporâneas do português brasileiro, devidamente focalizadas, podem fornecer evidências empíricas decisivas para a comprovação de hipóteses históricas. Tal assunção está na base da pesquisa sobre o português afro-brasileiro, cujos resultados estão reunidos neste livro.

Pode-se pensar, com base na reflexão de natureza historiográfica desenvolvida neste capítulo, que os reflexos do contato entre línguas se diluíram pelas diversas variedades do caleidoscópio sociolinguístico brasileiro, nas mais diversas configurações estruturais, em que os chamados fatores internos e externos se entrecruzam, por vezes de maneira inextricável. Em tal cenário, pode ser de grande significância isolar uma variedade linguística em que os reflexos de pretéritas situações de contato sejam mais visíveis. Originárias, em muitos casos, de antigos quilombos e mantidas em grande isolamento até meados do século XX, comunidades formadas em sua grande maioria por descendentes diretos de escravos africanos, como a comunidade de Helvécia no Extremo Sul do Estado da Bahia, podem se tornar verdadeiros sítios arqueológicos da história sociolinguística do Brasil. Formadas em situações de grande concentração de africanos com pouco acesso aos modelos da língua portuguesa no interior do país, essas variedades linguísticas podem ter sofrido uma erosão gramatical e passado por processos de reestruturação, em níveis superiores aos que afetaram outras variedades do português brasileiro. Por outro lado, o seu isolamento pode ter permitido que os reflexos desses processos de variação e mudança induzidos pelo contato entre línguas possam ter subsistido, não obstante a ação niveladora dos grandes centros urbanos, até as últimas décadas do século XX e princípios do século XXI, quando as amostras de fala do **Projeto *Vertentes do Português Popular do Estado da Bahia*** foram recolhidas.[55]

Em sua formulação mais simples e transparente, a hipótese reitora do programa de investigação de que resulta este livro é a seguinte: se o português brasileiro foi afetado em seu desenvolvimento histórico pelo contato entre línguas, os reflexos desse processo serão mais notáveis na fala das centenas de comunidades rurais afro-brasileiras que ainda subsistem no interior do país. O próximo capítulo tem por objetivo fundamentar essa hipótese, através de uma tentativa de formalização do conceito de **português afro-brasileiro**.

[55] A metodologia do **Projeto *Vertentes*** é descrita no capítulo 5 deste livro.

2 O português afro-brasileiro: as comunidades analisadas

Dante Lucchesi

Alan Baxter

Jorge Augusto Alves da Silva

*Cristina Figueiredo**

No panorama sociolinguístico do Brasil, as comunidades rurais afro-brasileiras isoladas ocupam uma posição própria, em virtude de certas especificidades que as caracterizam, do ponto de vista histórico, étnico e socioeconômico. Em sua maioria, essas comunidades se originaram em antigos quilombos ou em populações de escravos que receberam doações de terra, com o fim dos empreendimentos agroexportadores escravagistas. Desse modo, tais comunidades se definem pelos seguintes parâmetros: (i) são compostas majoritariamente por descendentes diretos de escravos africanos que se fixaram em localidades remotas do interior do país e de difícil acesso; (ii) mantiveram-se em relativo isolamento até a segunda metade do século XX; (iii) sua principal atividade econômica é a agricultura de subsistência.

A nativização do português entre os descendentes dos escravos africanos é determinante na história sociolinguística dessas comunidades. Os escravos vindos da África tiveram de aprender o português como segunda língua em condições bem adversas, no ambiente de trabalho forçado da lavoura ou das minas. Já os que nasciam no Brasil e, em muitos casos, eram filhos de pais que falavam línguas africanas diferentes e mutuamente ininteligíveis, adquiriam o português como língua materna a partir do modelo imperfeito de português falado como segunda língua pelos adultos. Esse processo de nativização de um modelo defectivo de segunda língua, aqui denominado de **transmissão linguística irregular**, teria dado origem a uma variedade linguística do português muito diferente do português falado pelos colonos portugueses e seus descendentes "brasileiros" (cf. capítulo 3

* A introdução do capítulo, bem como a seção 2.1 e a sua conclusão são de autoria de Dante Lucchesi, responsável também, juntamente com Alan Baxter, pela seção 2.2., relativa a Helvécia, e pela seção relativa aos arraiais de Rio de Contas, em co-autoria com Jorge Augusto Alves da Silva, autor da seção sobre Cinzento; a seção sobre Sapé é de autoria de Cristina Figueiredo.

deste livro). E, tendo em vista o isolamento em que essas comunidades se conservaram no interior do Brasil até pelo menos a primeira metade do século XX, essa variedade bastante alterada do português foi sendo passada de geração a geração sem maiores alterações, até sofrer a crescente influência do português urbano culto (ou semiculto) a partir das últimas décadas do século XX. Tal influência reflete-se nos padrões de variação e mudança linguísticas em curso hoje nessas comunidades, e que são analisadas na segunda parte deste livro.

Por serem marcadas em sua formação por situações de contato entre línguas maciço e abrupto, essas comunidades seriam um local privilegiado para a ocorrência de processos de crioulização do português. Contudo, não há registros históricos claros da existência de alguma comunidade crioulófona no Brasil. Talvez o mais próximo a que se tenha chegado disso sejam os registros feitos por Carlota Ferreira, no início da década de 1960, na comunidade afro-brasileira de Helvécia, situada no Extremo Sul do Estado da Bahia (FERREIRA, 1984). De seu breve contato com um casal de falantes idosos dessa comunidade, a pesquisadora concluiu pela existência de um antigo "falar crioulo", mas, por não dispor, na altura, de qualquer equipamento de gravação, os seus registros se resumem à transcrição fonética de alguns enunciados, e suas conclusões devem ser contextualizadas no estágio do conhecimento incipiente acerca das teorias sobre a crioulização no Brasil, na época.[1] Análises linguísticas feitas em outras regiões do país potencialmente crioulizantes, como o Vale da Ribeira, em São Paulo, também não chegaram a reconhecer algum "dialeto crioulo" do português (CARENO, 1991, p.105-106). Em suas pesquisas de campo durante a década de 1990, Alan Baxter e Dante Lucchesi entraram em contato com mais de dez comunidades rurais afro-brasileiras isoladas no interior dos Estados da Bahia, Espírito Santo e Piauí, e a comunidade com o comportamento linguístico mais desviante ainda foi Helvécia, sem se encontrar ali, cerca de trinta anos após o primeiro contato, uma variedade crioula do português (cf. seção 2.1 abaixo). Portanto, pode-se concluir que, se não há evidências empíricas claras acerca da ocorrência de processos de crioulização do português nas cerca de oitocentas comunidades rurais afro-brasileiras isoladas existentes no Brasil,[2] não há igualmente provas que refutem a possibilidade de que pelo menos algumas dessas comunidades tenham passado por processos de crioulização em sua história.

Por outro lado, o reflexo mais incontroverso do contato do português com as línguas africanas no comportamento linguístico das comunidades rurais afro-brasileiras parece ser o uso de certos códigos de base lexical africana, com o estatuto de língua secreta, que ainda se conserva em algumas dessas comunidades, como a de Cafundó, em São Paulo, e a de Tabatinga, em Minas Gerais (já referidas na

---

[1] A análise feita por Carlota Ferreira é descrita na seção 2.1.3 abaixo.

[2] Essa cifra é baseada na análise descritiva de Anjos (2000).

seção 1.1 do capítulo anterior). Mas, considerando o espectro funcional bastante restrito dessas línguas secretas, seu uso é de somenos importância para a formação dos padrões de comportamento linguístico dessas comunidades no cenário dos falares rurais brasileiros, não obstante a sua grande relevância no plano simbólico da construção da identidade e da resistência cultural.

Diante disso, mantém-se a visão desenvolvida no capítulo anterior de que as variedades populares do português no Brasil, sobretudo no interior do país (onde vivia a maioria da população até a primeira metade do século XX), se formaram em situações de transmissão linguística irregular de tipo leve. E, nesse universo dos falares rurais brasileiros, as comunidades isoladas de afrodescedentes seriam caracterizadas pela ocorrência de um processo de alteração um pouco mais profundo, podendo, em alguns casos excepcionais, aproximarem-se da crioulização.

A eventual excepcionalidade do caso de Helvécia tem sido atribuída à presença de suíços, franceses a alemães à frente dos empreendimentos agroexportadores que deram origem à comunidade (ZIMMERMANN, 1999); mas a questão é controversa (cf. seção 2.1.2, abaixo). A identificação de um estágio mais crioulizado do dialeto de Helvécia pode estar relacionada à época em que foi feito o primeiro contato, pois nenhuma das comunidades aqui analisadas foi visitada por algum linguista antes da década de 1990. Pode-se argumentar que, nas primeiras décadas do século XX, várias comunidades rurais afro-brasileiras isoladas poderiam ainda conservar marcas mais evidentes de um estágio anterior de crioulização, como as que, segundo Ferreira, ainda podiam ser observadas na fala dos membros mais idosos da comunidade de Helvécia, no início da década de 1960. Com efeito, se as análises apresentadas neste livro identificam processos linguísticos exclusivos da comunidade de Helvécia no conjunto de comunidades rurais afro-brasileiras aqui observadas (como é o caso da variação na concordância gênero no interior do Sintagma Nominal e a variação na concordância verbal junto à primeira pessoa do singular), o comportamento linguístico dessa comunidade não difere muito do das demais na maioria dos processos de variação aqui descritos. Tais fatos revelam a complexidade da tarefa de identificar diferenças no comportamento linguístico de comunidades de fala similares a partir de especificidades de sua formação histórica. Nesse plano, coloca-se a questão de compreender os processos de formação histórica da variedade linguística das comunidades rurais afro-brasileiras isoladas em cotejo com o que se observou em outras regiões do continente americano em que também houve uma grande concentração de africanos, em decorrência do tráfico negreiro.

O sequestro de grandes contingentes de indivíduos do continente africano para serem empregados no trabalho forçado em empresas agroexportadoras constitui um dos elementos centrais do processo de colonização europeia do continente americano, na região do Caribe, do Sul dos EUA e do Brasil. A aquisição da língua do grupo dominante por parte desses africanos escravizados deu-se em situações

tão precárias, que deu ensejo a novas línguas, denominadas *crioulas*, as quais, não obstante o fato de seu léxico ser, em sua grande maioria, derivado da língua dominante (com notáveis alterações fonéticas), a sua gramática é qualitativamente distinta da dessa língua que lhe forneceu o grosso do vocabulário. Desse modo, o Caribe é uma das regiões do planeta que concentra um grande número das línguas crioulas hoje conhecidas, de diferentes bases lexicais (francês, inglês, holandês, etc.). Também se registram casos incontestes de crioulização do inglês em comunidades isoladas de afrodescendentes no Sul dos Estados Unidos (HOLM, 2004). Apesar de ter havido um processo semelhante de colonização no Brasil, a complexidade da sociedade brasileira, com suas especificidades (sobretudo o seu elevado grau de mestiçagem), não possibilitou a ocorrência de um processo representativo e duradouro de crioulização do português. Mas não se pode deixar de pensar que as variedades do português que se formaram na grande massa de afrodescendentes, indiodescendentes e mestiços, que formam a base da sociedade brasileira (em diferentes proporções, consoante a região do país), passaram por processos de variação e mudança semelhantes aos que estão na base da formação das línguas crioulas, contudo em um grau de intensidade menor (cf. discussão do capítulo anterior).

O pressuposto de que as variedades populares do português brasileiro foram afetadas por mudanças induzidas pelo contato entre línguas em sua formação está na base da proposição de uma variedade *afro* do português brasileiro. O mais provável é que as variedades do português que se formaram em torno de *plantations*, como os grandes engenhos do Nordeste, ou em agrupamentos de escravos foragidos, tenham sido as que mais foram afetadas pelo contato. É claro que o contato também deixou seus efeitos em outros contextos sociolinguísticos, como o da grande massa de escravos urbanos nas cidades, na mineração (sobretudo no século XVIII) ou mesmo na pecuária, conquanto não sejam esses contextos propícios à crioulização. Portanto, tem-se um quadro matizado, com matizes mais fortes de um lado, em que os efeitos do contato são mais profundos, até o outro extremo, em que os seus efeitos são mais superficiais (pode-se pensar, nesse caso, nos segmentos da população branca, de falantes nativos do português, que conviveram mais intimamente com a escravaria). Esse quadro torna-se ainda mais complexo e imbricado, considerando-se o tráfico interno de escravos (como a venda em massa dos escravos dos engenhos do Nordeste para a região das minas, no século XVIII) e os imensos fluxos migratórios (como o grande êxodo dos ex-cativos para a periferia das cidades, após a abolição).[3]

Mas, apesar de toda essa imbricação dos possíveis efeitos do contato entre línguas nas variedades atuais do português brasileiro, é lícito deslindar, com base na fala das comunidades rurais afro-brasileiras isoladas, a variedade do português

[3] Cf. Mattoso (2003).

no Brasil que mais foi afetada pelo contato entre línguas. Isso implica que o português afro-brasileiro seria diferenciado mesmo em face das demais variedades populares e rurais do país. A diferenciação das comunidades rurais afro-brasileiras isoladas frente às demais comunidades de fala rurais brasileiras, em termos linguísticos, está consubstanciada na identificação de processos linguísticos presentes nas primeiras e ausentes nas últimas, bem como em uma maior intensidade dos processos de variação nas primeiras em relação às demais. E, dentro de uma concepção dialética da produção do conhecimento, ao tempo em que essas descobertas empíricas possibilitam a identificação dos processos de variação e mudança desencadeados em situações de contato entre línguas, tais evidências empíricas também podem ser tomadas como comprovação da hipótese de que o contato entre línguas foi relevante na conformação dessas variedades linguísticas. Tal processo, entretanto, não se encerra nesse circuito empírico, na medida em que são mobilizadas as teorias desenvolvidas independentemente a partir de observações de outras situações de contato linguístico, nomeadamente aquelas que deram ensejo à formação de línguas reconhecidamente crioulas; ou seja, os parâmetros que se adotam aqui para a identificação dos processos de variação e mudança induzidos pelo contato entre línguas na fala das comunidades rurais afro-brasileiras isoladas são, sempre que possível, aqueles que a crioulística tem relacionado com as situações em que os efeitos do contato entre línguas são evidentes. Mas, como se pode ver no capítulo 3 deste livro, ainda não há consenso sobre essas questões na crioulística, assim como não os há em outras áreas da ciência da linguagem; o que não constitui novidade alguma, desde que não se adote uma visão *naïve* de ciência.

Essa possibilidade de traçar fronteiras entre variedades linguísticas definida nesse plano de investigação não deve ser tomada como algo simples e cristalino. A linguística tem se defrontado, sem progressos retumbantes, com a tarefa de traçar fronteiras entre línguas e/ou dialetos a partir de variedades linguísticas diferenciadas, nos planos histórico, político e cultural. As dificuldades são ainda maiores quando se trata de deslindar variedades linguísticas em uma população reunida historicamente em uma mesma unidade administrativa e que se agrega em torno de uma mesma identidade cultural de nação, compartilhando também o mesmo sistema de educação formal, bem como os meios de comunicação de massa, e sobretudo que se identifica como uma comunidade linguística única, na medida em que se adota a mesma norma linguística de referência.[4] Mas, apesar de sua

[4] Toda essa identidade que aqui se atribui à população brasileira, longe de ser homogênea, deve ser, obviamente, matizada em função das diferenças sociais, de inserção no processo produtivo, de apropriação da riqueza, de relações de dominação e exclusão, de produção simbólica, em suma: em função das relações entre as classes sociais dentro do sistema de produção capitalista.

complexidade, a tarefa cumpre um objetivo crucial para o avanço do conhecimento acerca da realidade sociolinguística do Brasil, bem como de sua formação histórica, na medida em que pode estabelecer correlações empiricamente motivadas entre processos sócio-históricos e processos de estruturação linguística.

Assim, a identificação de processos de variação derivados de mudanças induzidas pelo contato entre línguas em comunidades rurais afro-brasileiras, que não se encontram em outras comunidades rurais brasileiras não marcadas etnicamente, ou uma maior intensidade em tais processos nas primeiras em relação às últimas, constitui uma evidência empírica significativa de que o contato entre línguas afetou o português brasileiro em condições históricas bem determinadas. As condições históricas que caracterizam a formação das comunidades afro-brasileiras, conforme descritas acima, fizeram com que os efeitos do contato entre línguas, não apenas fossem mais intensos na formação da gramática dessas comunidades de fala, como se conservassem com mais nitidez até os dias atuais. Já em uma comunidade rural formada com uma maior proporção de colonos mestiços, ou mesmo de descendentes de colonos portugueses, seriam afetados pelo contato entre línguas com menor intensidade.

Os avanços do que se pode chamar de linguística sócio-histórica em estabelecer correlações empiricamente motivadas entre processos sócio-históricos e processos de estruturação linguística são, até os dias de hoje, bem modestos, por várias razões. Em primeiro lugar, pela histórica hegemonia da concepção imanentista de língua em toda ciência da linguagem, mesmo em suas correntes historicistas, de modo que, no geral, a história de uma variedade linguística é sempre analisada nos termos da lógica de sua, muitas vezes hipostasiada, estrutura interna. Esse estado de coisas coaduna-se perfeitamente com a dificuldade inerente à tarefa de estabelecer relações empiricamente motivadas em universos de observação mais complexos, como no plano das organizações sociais e sua história. Essa dificuldade se traduz, no plano da investigação científica, nas limitações relativas à observação controlada dos fatos, associada à incapacidade de gerar modelos de análise que possam integrar um conjunto potencialmente infinito de variáveis, que deveriam ser consideradas para explicitar a determinação global dos processos estudados. Essa complexidade em estabelecer determinações, com controle empírico, no plano social e da cultura, é particularmente crítica quando se pensa no caráter onipresente da linguagem verbal no plano das relações humanas. Por outro lado, ergue-se, como grande barreira para a observação empírica, a desoladora escassez de registros históricos de variedades que não sejam as padronizadas para o uso oficial e erudito, em função da poderosa apropriação política e ideológica dos usos linguísticos em todas as sociedades de classe. Pode-se, portanto, entender por que a ciência da linguagem avançou tão pouco em estabelecer relações empiricamente motivadas entre os processos sócio-históricos e processos de estruturação

linguística, não indo muito além, até hoje, do estabelecimento de relações, no mais das vezes, mecanicistas e de valor heurístico questionável entre fatores sociais e variáveis linguísticas, a partir de abordagens sincrônicas.[5]

No que concerne ao objeto de estudo deste livro, essas adversidades que se colocam para a pesquisa empírica são ainda mais inclementes, na medida em que está se tratando de comunidades historicamente exploradas, oprimidas e marginalizadas. Os registros históricos, na maioria dos casos, inexistem. Mesmo no plano da tradição oral, elementos fundadores dessas comunidades como a escravidão ou a fuga quilombola são recalcados, seja como mecanismo de defesa, seja por assimilação da ideologia dominante. Além disso, a complexidade e a diversidade das redes de determinações histórico-sociais sobre os processos linguísticos, acima referidas, faz com que comunidades de fala que compartilham processos semelhantes de formação, no plano histórico-social, como é caso das comunidades afro-brasileiras, não apresentem hoje necessariamente os mesmos padrões coletivos de comportamento linguístico.

A intensidade com a qual o contato entre línguas afetou essas comunidades em sua formação variou bastante em função da conjugação particular, em cada caso, de uma série de fatores históricos, tais como: o grau de acesso dos antepassados africanos aos modelos de português disponíveis, a proporção entre africanos e crioulos decorrente da taxa de natalidade no grupo, o grau de heterogeneidade linguística entre os africanos, entre tantos outros (cf. seção 3.2 do próximo capítulo). Por outro lado, a conservação desses processos até os dias de hoje dependerá da manutenção do isolamento da comunidade, que tem se reduzido de forma crescente nas últimas décadas. Disso decorre, obviamente, que haverá comunidades afro-brasileiras cujos padrões de comportamento linguístico exibirão mais reflexos de processos de variação e mudança resultantes do contato entre línguas do que outras. Diante disso, é necessária a adoção de um critério não linguístico, isto é, independente, para definir o que seja uma comunidade rural afro-brasileira, como se fez no primeiro parágrafo deste capítulo, com base no protótipo de uma comunidade rural isolada formada majoritariamente por afrodescendentes que se dedicam precipuamente à agricultura de subsistência.

Não se podendo, portanto, pensar o **português afro-brasileiro** como uma realidade linguisticamente homogênea, a diferença entre ele e o que se pode chamar de **português rural brasileiro**, ou mesmo **português popular do interior do país** será igualmente variável em função do recorte feito, ou seja, das comunidades em cotejo em cada caso. E, mais uma vez, impõe-se o esquema de um *continuum*. Em um extremo, estão aquelas comunidades rurais afro-brasileiras isoladas mais

---

[5] Para uma discussão dessa questão, vejam-se, entre outros: Lass (1980), Lucchesi (2004c) e Faraco (2005).

afetadas pelo contato entre línguas, em cuja formação se pode pensar que teriam ocorrido processos muito próximos ao que se reconhece hoje em dia como *crioulização*. A partir desses poucos casos, vem um crescente número de comunidades em cuja formação predominam processos cada vez mais leves de transmissão linguística irregular. Quanto maior a proporção de afrodescendentes diretos na comunidade, maior será o grau de interferência da transmissão linguística irregular na formação daquele dialeto, o que não significa necessariamente um processo mais intenso, uma vez que isso, como já se disse acima, é uma resultante da combinação de uma série de fatores históricos. De qualquer forma, e prosseguindo no *continuum*, chega-se às comunidades rurais mais nitidamente mistas, com um crescente contingente de mestiços e de brancos, até alcançar as comunidades com um percentual reduzido, ou mesmo ínfimo de afrodescendentes (ou indiodescendentes, dependendo da região). Essas últimas seriam as que menos teriam sido afetadas pelo contato entre línguas em sua formação, podendo-se mesmo pensar, em alguns poucos casos, apenas em uma influência indireta por meio do **contato dialetal**.

Portanto, pode-se concluir em relação ao universo das comunidades rurais brasileiras que a composição étnica atual tem implicações históricas potenciais no seguinte sentido: quanto maior a proporção de afrodescendentes (ou indiodescendentes, consoante a região), maior será a probabilidade de a gramática daquela comunidade de fala ter sido afetada pelo processo de transmissão linguística irregular desencadeado em situações de contato linguístico massivo (sendo a intensidade de tal processo, em cada caso, mediada por uma série de fatores sócio-históricos). Define-se, assim, um *continuum* em um universo em que predominam comunidades mistas, de modo que, considerando-se o contexto histórico de formação de todas as comunidades rurais brasileiras, é difícil conceber alguma que não tenha sido afetada pelo contato entre línguas em maior ou menor grau.[6]

Esse quadro define a importância da descrição sistemática da gramática das comunidades de fala afro-brasileiras para a compreensão da história sociolinguística do país. Os processos linguísticos que ainda podem ser observados nessas comunidades podem refletir processos de variação e mudança muito representativos na história linguística do Brasil, sobretudo entre os séculos XVII e XIX; período em que o português tinha de conviver com centenas de línguas indígenas e africanas. E considerando que, desde então, os falantes nativos do português europeu e seus descendentes diretos constituíam apenas um terço de toda a

[6] Deve-se considerar também as comunidades rurais ligadas à imigração europeia e asiática, ocorrida entre a segunda metade do século XIX e a primeira metade do século XX, mas com implicações linguísticas diferenciadas, em função das diferenças significativas entre a fixação desses segmentos no campo brasileiro *vis-à-vis* aos africanos e seus descendentes.

população, esses processos de variação e mudança induzidos pelo contato entre línguas pode estar na base de tudo isso que se denomina **português brasileiro**.

Dentro dessa perspectiva, este capítulo contém uma breve descrição sociodemográfica e histórica das quatro comunidades rurais afro-brasileiras isoladas que constituem o universo de observação das análises contidas neste livro, estruturada da seguinte maneira: na seção 2.1 é feita uma apresentação geral das quatro comunidades; a seção 2.2 descreve a comunidade de Helvécia; a seção 2.3 é dedicada á comunidade de Cinzento; a seção 2.4, aos arraiais de Rio de Contas, e a seção 2.5, à comunidade de Sapé. Uma breve conclusão encerra o capítulo.

2.1 As comunidades de fala analisadas

A importância de uma descrição sistemática da gramática das comunidades rurais afro-brasileiras isoladas para a compreensão do processo histórico de formação das variedades da língua portuguesa do Brasil, tanto no sentido de identificar que mudanças foram desencadeadas pelo contato entre línguas, quanto para estabelecer relações empiricamente motivadas que atestem que o contato entre línguas constituiu condição necessária e suficiente para que tais mudanças fossem desencadeadas, impõe a necessidade de que tal descrição seja feita sobre uma sólida base empírica. O conjunto de análises de tópicos da morfossintaxe que compõem a segunda parte deste livro tem como base empírica o **Acervo de Fala Vernácula do Português Afro-Brasileiro do Estado da Bahia**, constituído no âmbito do **Projeto** ***Vertentes do Português Popular do Estado da Bahia*** (<http://www.vertentes.ufba.br>). Esse acervo reúne amostras de fala vernácula de quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia.[7]

Considerando a heterogeneidade relativa do português afro-brasileiro (conforme argumentação acima), foi preciso fazer uma seleção de comunidades rurais que fosse minimamente representativa do universo de mais de duzentas comunidades rurais afro-brasileiras que, segundo Anjos (2000), existem no Estado da Bahia. Se as análises feitas se baseassem apenas nos padrões de fala de uma única comunidade, como, por exemplo, Helvécia, poderia ser argumentado que os achados dessa investigação seriam excepcionais e representativos apenas daquela comunidade estudada. Portanto, na seleção de comunidades que integrariam o *corpus* do português afro-brasileiro, buscou-se uma representatividade do ponto de vista quantitativo e, principalmente, do ponto vista qualitativo. Isso implicou não apenas a escolha de várias comunidades, como também a escolha de comunidades representativas de contextos históricos e socioeconômicos diversos.

---

[7] A metodologia empregada na constituição dessas amostras é descrita no capítulo 5 deste livro.

Com base nesses parâmetros de representatividade, foram escolhidas, como locais de recolha de fala vernácula do português afro-brasileiro, quatro comunidades de diferentes regiões do Estado da Bahia: Helvécia, no Município de Nova Viçosa; Cinzento, no Município de Planalto; as comunidades geminadas de Barra e Bananal, no Município de Rio de Contas; e Sapé, no Município de Valença. Helvécia, situada no Extremo Sul, perto do litoral, tem a sua origem associada a empresas de cultivo e exportação de café, estabelecidas na região, no século XIX, por colonos suíços, franceses e alemães, com base no largo emprego de mão de obra escrava trazida da África. Após a abolição e o ocaso do empreendimento agroexportador, a grande maioria dos ex-escravos manteve-se na região, dedicando-se à agricultura de subsistência nas terras abandonadas pelos colonos estrangeiros. Cinzento se originou provavelmente a partir de um agrupamento de escravos foragidos em uma remota localidade no Semiárido baiano, numa região tradicionalmente ligada à pecuária e que tem hoje como grande centro urbano a cidade de Vitória da Conquista. As comunidades de Barra e Bananal, tradicionalmente chamadas de arraiais de Rio de Contas, situam-se nesse município da Chapada Diamantina, cujo povoamento remonta ao final do século XVII e tem sua história intimamente ligada à mineração. Após a pujança do século XVIII, veio a decadência com o esgotamento dos veios auríferos, e só nas últimas décadas a cidade de Rio de Contas voltou a aquecer a sua economia com base na exploração do turismo. É incerto o caráter original dos arraiais, não estando descartada a sua natureza quilombola. A comunidade de Sapé está situada ao sul do Recôncavo Baiano, região historicamente marcada pelo estabelecimento de inúmeros engenhos de açúcar, já a partir da segunda metade do século XVI, bem como por plantações de fumo, sobretudo a partir do século XVIII. Circundando a cidade do Salvador, no fundo da Baía de Todos os Santos, essa região manteve uma estreita relação com a capital, com um intenso comércio que se viabilizava através de um sem número de saveiros que cruzavam a grande baía até as primeiras décadas do século XX. A comunidade se formou, provavelmente, por meio da ocupação de terras áridas e de difícil acesso, “doadas” por senhores decadentes a ex-escravos.

Dentre essas comunidades, a que mais tem atraído a atenção dos linguistas é Helvécia, por conta dos registros feitos por Carlota Ferreira no início da década de 1960.[8] O maior volume de informações sobre a comunidade vai se refletir na descrição de aspectos históricos e socioeconômicos das quatro comunidades, que se faz nas restantes seções deste capítulo. A descrição de Helvécia é a mais ampla de todas. Não obstante a virtual importância de Helvécia na identificação de

[8] Entre os estudos linguísticos que se fizeram desde então sobre Helvécia, podem ser citados: Baxter (1992, 1996, 1998, 1999), Holm (1992), Megenney (1993), Baxter e Lucchesi (1993 e 1997), Baxter, Lucchesi e Guimarães (1997) e Zimmermann (1999).

processos de variação e mudança mais radicais, desencadeados pelo contato entre línguas, um estudo coordenado de quatro comunidades diferentes é importante até para aferir a heterogeneidade do português afro-brasileiro. Por outro lado, essa representatividade possibilita que as análises aqui apresentadas possam fundamentar a constituição da identidade linguística do português afro-brasileiro, na medida em que a comparação com outras variedades do português brasileiro bem como o estabelecimento de paralelos com o que se observa em línguas crioulas de base lexical portuguesa sejam feitos a partir de uma sólida base empírica.

2.2 A comunidade de Helvécia

Em 1818, o Governo Central concedeu a três colonos alemães sesmarias ao longo das margens do Rio Peruípe, doze léguas acima de Nova Viçosa, na Comarca de Caravelas, no Extremo Sul da Província da Bahia. Logo em seguida, juntamente com outros colonos suíços e franceses, eles fundaram a Colônia Leopoldina, um conjunto de plantações que, utilizando o trabalho escravo, se dedicava ao cultivo e à exportação do café. Foi um dos últimos empreendimentos agrícolas no Brasil a utilizar o trabalho escravo em larga escala, num tempo em que a instituição da escravidão sofria uma forte pressão externa. Mas a colônia rapidamente prosperou e, no início da década de 1850, era responsável por quase 90% da produção de café da província.

Entretanto, uma combinação de fatores conspirou para a sua rápida decadência, sobretudo o advento da abolição e a expansão da cultura do café para regiões mais férteis da província, de modo que, no final da década de 1880, o Recôncavo já havia se tornado o maior centro cafeeiro da Bahia (BAHIA, 1978). Nessa altura, grande parte dos colonos já havia deixado a região, muitos dos quais retornando à Europa. Porém, após a abolição, muitos dos ex-cativos permaneceram nas terras vizinhas às antigas plantações, praticando uma cultura de subsistência; no que foram sucedidos por seus descendentes. Já a vila de Helvécia passaria a viver em função da estação da ferrovia Bahia-Minas, inaugurada em 1897.

Atualmente, a única marca da antiga presença dos antigos colonizadores europeus é o nome da localidade, estampado no prédio da agora desativada estação ferroviária, localizada no centro da vila que cresceu nas terras da antiga plantação "Helvécia", do suíço Johannes Martinus Flach. Por outro lado, a grande concentração de escravos ocorrida no século passado ainda se faz notar na composição étnica da população local. De resto, a comunidade de Helvécia está normalmente inserida na economia da região do Extremo Sul do Estado da Bahia.

2.2.1 Caracterização socioeconômica e demográfica

O Distrito de Helvécia está situado a 17º 48' 27" S de latitude e 39º 39' 51" W de longitude, no Município de Nova Viçosa, na microrregião de Porto Seguro, na mesorregião do Sul baiano. A sua população, em 1996, perfazia um total de 16.474 habitantes, dos quais 15.108 viviam na zona rural, e apenas 1.366 constituíam a população da pequena vila de Helvécia. Do total dos seus 13.310 habitantes com mais de cinco anos, 7.203 eram alfabetizados; o que corresponde a uma taxa de alfabetização de 54% que prevalece tanto na zona rural (com 6.484 indivíduos alfabetizados, em um total de 11.972), quanto na zona urbana (719 indivíduos, em um total de 1.338). Na faixa da população com mais de sessenta anos, essa taxa de alfabetização cai para 22% (195 em um total de 887) (IBGE, 1996).

No centro da vila, encontra-se o prédio da há muito desativada estação ferroviária da antiga estrada de ferro Bahia-Minas. A estação foi construída em 1897 e ficava a 73 quilômetros do ponto zero da ferrovia, a estação de Ponta de Areia, situada no litoral baiano.[9]

A economia da região é fundamentalmente baseada na agricultura, na pecuária e na indústria de celulose. Recentemente, o turismo balneário tem se desenvolvido, sobretudo na sede do Município de Nova Viçosa, situada na costa. A maioria dos informantes que compõem a amostra de fala recolhida na comunidade trabalha em pequenas propriedades agrícolas familiares, dedicando-se sobretudo à cultura de subsistência. O pequeno e eventual excedente é vendido nas feiras de Helvécia ou de Teixeira de Freitas. Esse excedente provém normalmente do fabrico artesanal da farinha de mandioca. Cultivam também feijão, milho, arroz, abóbora, batata e outros legumes. Os mais jovens e mais prósperos chegam a empregar alguma maquinaria e adubos e fertilizantes químicos no cultivo de hortifrutigranjeiros, que são absorvidos pelo comércio da região. Muitos dos mais antigos trabalharam na construção e no funcionamento da ferrovia Bahia-Minas. Os mais novos empregam-se eventualmente nas plantações de eucalipto, de onde é extraída a celulose. A implantação dessas grandes plantações de eucalipto alterou a fisionomia da região, desalojando muitas famílias que viviam do cultivo em suas pequenas propriedades. Esses indivíduos e seus descendentes, agora, trabalham na terra de parentes, ou se empregaram em grandes propriedades agrícolas. Muitos se dirigem, ciclicamente, ao mercado de trabalho urbano, como vigilantes, operários

[9] De acordo com a informação prestada pela Rede Ferroviária Federal, com sede no Bairro da Calçada, em Salvador-BA, a ferrovia Bahia-Minas foi construída por partes, tendo sido iniciada a construção em 1853, e a última parte foi concluída em 1942. A ferrovia era administrada pela Viação Férrea Federal Leste Brasileiro.

da construção civil, empregadas domésticas, etc., tanto nos centros urbanos regionais, como Teixeira de Freitas-BA e Nanuque-MG, quanto nos grandes centros do Sul do país.

O acesso ao Distrito de Helvécia, bem como à sede do Município de Nova Viçosa, era feito, em 1994, por uma estrada vicinal de terra batida, a partir da BR-101. O município está na área de influência do centro regional de Teixeira de Freitas. Entretanto, alguns aposentados ainda preferem receber o seu provento na sede do Município de Caravelas. O hábito de ouvir rádio e ver televisão tem se difundido bastante, principalmente entre os mais jovens. Algumas famílias já possuem seu próprio aparelho de televisão; os demais contam com o aparelho que é ligado na praça pública, ao cair da tarde.

2.2.2 Sócio-história

Como já foi dito, a comunidade de Helvécia tem origem na antiga Colônia Leopoldina, uma colônia suíço-alemã estabelecida em 1818 e que prosperou até a abolição, em 1888, dependendo da mão de obra escrava para o cultivo do café. Segundo testemunho do médico da colônia, em 1858, havia uma população de 200 brancos, principalmente suíços e alemães, com alguns franceses e brasileiros e, 2.000 negros, na maior parte nascidos na colônia (TÖELSNER, s.d. apud FERREIRA, 1984, p.22). Alguns autores têm comentado que o exemplo de Helvécia é especial, porque os senhores eram, na sua maioria, estrangeiros e, portanto, não tinham o português como língua materna.[10] Contudo, há outros fatores mais significativos que apontam para o seu estatuto especial. Dentre esses, destaca-se o fato de a comunidade de ex-escravos da Colônia Leopoldina ter se fixado na região da colônia, numa situação relativamente isolada, e não se ter dispersado tanto quanto outras populações de ex-escravos, após 1888.

Por outro lado, ainda é possível a identificação dos fatores extralinguísticos que teriam determinado as profundas mudanças que marcaram a origem desse dialeto, graças a uma documentação histórica rica, embora incompleta, que facilita um estudo da demografia das populações de escravos e que permite formular algumas hipóteses em torno da questão.

Relativamente ao período anterior a 1850, o ano da Lei Eusébio de Queiroz que proibiu definitivamente a importação de escravos, há informações pertinentes à proporção de escravos para livres. Assim, em 1848, esta proporção era de 10 para 1, embora a proporção nas fazendas com grande número de escravos fosse bastante superior. No estabelecimento do suíço Johannes Martinus Flach, por exemplo, a

---

[10] Cf. discussão em Zimmermann (1999) e Baxter e Lucchesi (1999).

proporção era de 24 para 1, e havia 108 escravos (BAXTER; LUCCHESI, 1999, p.128-129). A julgar pelas evidências disponíveis, no período anterior a 1830, a situação seria semelhante (BAXTER, 1999, p.3-4).

É evidente que, em tais circunstâncias, o acesso que a criança escrava teria a modelos de falantes nativos do português seria mínimo. Mesmo o contato com o português falado como segunda língua, dos senhores estrangeiros, teria sido precário nas fazendas, em função do elevado número de escravos. Também cabe observar que a grande maioria dos escravos dessas fazendas eram escravos de lavoura. As listas de escravos de 1854 a 1882 manifestam uma média de 92% de escravos de lavoura, de maneira que supomos que os contatos que o escravo teria com modelos de português falado como língua materna seriam mínimos. Os modelos dominantes para a aquisição da língua materna da criança escrava seriam os modelos falados por outros escravos. Portanto, cabe investigar as proporções relativas de escravos africanos e escravos brasileiros adultos para esclarecer as proporções relativas de falantes de português como língua materna e de falantes de português como língua segunda, na população escrava.

Nesse mesmo período, e até o final da década de 1850, a proporção de africanos para crioulos nas fazendas ainda era bastante alta, atingindo uma média próxima a 50% da população adulta. Em determinadas fazendas, a proporção de africanos na população adulta era relativamente alta, ainda na década de 1860, como na fazenda da família Vequet, onde, em 1865, 59% dos 46 escravos adultos eram africanos, e na fazenda do médico Töelsner, onde, em 1864, a proporção era de 67% (16 em um total de 24 escravos). Mesmo na década de 1870, havia fazendas com proporções altas para aquela época. Um exemplo seria, em 1871, a fazenda da família Tatet, onde havia 43% de adultos africanos (24 em um total de 36); e, em 1872, a fazenda de Ana Ida, onde havia uma proporção de 37% (23 em 62). A Figura 1 apresenta as proporções para o período de 1850 a 1880:

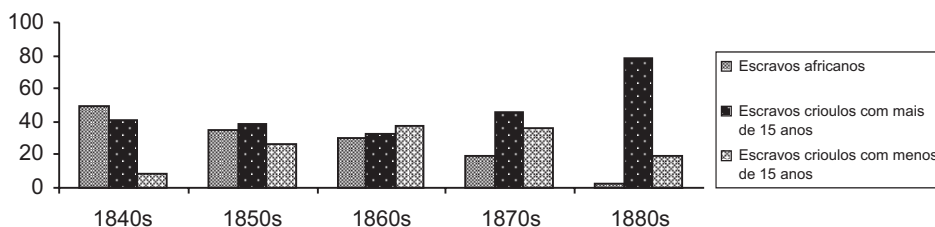



Figura 1 - Proporções de escravos africanos e de escravos brasileiros adultos e pré-adolescentes nas fazendas da Colônia Leopoldina

Fonte: Baxter e Lucchesi, 1999.

As implicações dessas proporções altas nos primeiros períodos são bastante óbvias. O elemento africano teria fornecido para as crianças escravas nascidas na colônia modelos muito defectivos de português falado como segunda língua (e,

talvez, em algumas circunstâncias, o modelo seria mesmo o das línguas africanas). Muito importante nesse sentido é o fato de que havia uma população crescente de escravos nascidos na Colônia Leopoldina. Havia, mesmo, um alto grau de filiação e continuidade geracional nas populações escravas em muitas das fazendas da Colônia Leopoldina (BAXTER, 1999, p.17-21). Em outras palavras, parece que houve uma política de criar escravos. Seja como for, criaram-se populações bastante estáveis em contextos herméticos; fatores que teriam contribuído para a manutenção das caraterísticas da(s) variedade(s) de português desenvolvida(s) entre os escravos que iam nascendo na colônia.[11]

Contudo, no que diz respeito à natureza dos modelos disponíveis para o processo de nativização das variedades de português falado como língua segunda entre os escravos, é evidente que são os contextos específicos que são pertinentes. Uma das listas de escravos consultadas permite uma visão muito ilustrativa dos possíveis modelos disponíveis à criança numa população escrava duma fazenda de café, no que diz respeito às línguas africanas e aos modelos de português falado pelos escravos adultos. A lista provém do inventário Mantandon, de 1858 (BAXTER, 1999, p.12-14), e apresenta a seguinte composição da escravaria: 23 africanos, 33 crioulos e uma mulata. Entre os escravos nascidos no Brasil, há 18 crianças pré-adolescentes. Os africanos são de seis grupos etnolinguísticos diferentes — monjolo (1), nagô (4), jeje (1), cabinda (1), moçambique (1), benguela (2) —, e há oito africanos de origem não determinada. De especial interesse é o fato de os casais e os conjuntos familiares estarem identificados. Há cinco casais africanos, um dos quais (pai cabinda e mãe moçambique) tem filhos. Há dois casais africano/crioulo (pai crioulo e mãe benguela), os dois com filhos; e há três casais crioulos, dois deles com filhos. E há ainda seis mães solteiras.

Quais eram os modelos linguísticos potencialmente disponíveis para as crianças nesse contexto? Algumas ilações podem ser feitas, assumindo os pressupostos de que: (i) o português já se teria nativizado entre os escravos nascidos no Brasil; e (ii) a fala da mãe constitui o modelo predominante no processo de aquisição da língua materna. Por um lado, as crianças teriam o português falado como língua materna por escravos nascidos no Brasil. De fato, oito grupos de crianças tinham mães crioulas. Desses oito grupos de irmãos, dois tinham também o pai nascido no Brasil. Mas não é possível precisar o grau de alterações presente na(s) variedade(s) de português falada(s) por esses crioulos, podendo-se mesmo pensar que esses indivíduos já falassem uma variedade crioulizada de português decorrente da nativização de variedades bem defectivas de português falado como segunda língua (cf. capítulo 3 deste livro). Por outro lado, há três grupos de

[11] As implicações linguísticas de processos desse tipo são discutidas no próximo capítulo.

crianças com modelos maternos de português falado como segunda língua; num desses casos, o pai também não é falante nativo do português. Além disso, existe a possibilidade de que as crianças de três unidades familiares tenham adquirido também a língua africana dos seus pais. Para a criança, nessa fazenda, fora do contato dos pais, as possibilidades de contato com modelos de português falado como segunda língua pelos outros escravos são, em termos puramente numéricos, maiores do que as possibilidades de contatos com o português nativizado falado pelos escravos. O fato de que 55% dos escravos adultos eram africanos também implica que a criança entraria em contato com línguas africanas, talvez na senzala. Pelo menos sete origens etnolinguísticas africanas estão representadas nessa população, e é provável que pelo menos três línguas africanas fossem faladas: nagô, congo e benguela.[12]

Um outro aspecto importante desse gênero de microssituação diz respeito às fontes de uma eventual influência do substrato. Nesse caso, identificaram-se sete possíveis fontes de influências do substrato, com mais probabilidade para as línguas com mais falantes. Contudo, noutras fazendas, a gama de fontes dessas influências podia ser ainda mais extensa, como foi o caso da Fazenda Krull, em 1854, onde havia 33 africanos com as seguintes caraterísticas: origem não identificada (4), nagô (14), cabinda (5), congo (3), hauçá (2), benim (1), calabar (1), moçambique (1), rebola (1), jeje (1) (BAXTER; LUCCHESI, 1999, p.131). Embora os escravos do grupo linguístico kwa predominassem em algumas fazendas, em outras fazendas preponderavam escravos do grupo linguístico banto. Na fazenda da família Reis, em 1854, só havia escravos do grupo linguístico banto, e eram de seis áreas linguísticas, sendo predominantes os moçambiques (BAXTER; LUCCHESI, 1999). Essa diversidade interfazenda e intrafazenda implica que as influências do substrato tenderiam a ser diluídas na maioria das situações e muito específicas em uns poucos casos.

Para o período de 1847 a 1872, os números totais dos africanos identificados por origem indicam que, proporcionalmente, houve mais escravos do grupo banto, de diversas áreas, porém os nagôs constituíram o maior grupo uniforme (BAXTER, 1999, p.7-8). Nesse sentido, além da possibilidade da prevalência de algumas estruturas do substrato no processo de aquisição/nativização do português, deve-se pensar que seja muito provável a utilização de línguas africanas em determinadas fazendas durante as primeiras décadas da colônia. Contudo, os efeitos dessa presença linguística africana seriam diluídos a partir de 1850, com o fim da importação de africanos, o envelhecimento da população africana e a mistura de escravos por meio do comércio interno à própria colônia (BAXTER, 1999).

---

[12] Adota-se a suposição de que as origens anotadas para os escravos nessas listas podem servir como indicadores da identidade etnolinguística. Dessa maneira, os escravos anotados como "congo" e "benguela" foram tratados como sendo efetivamente dessas áreas linguísticas banto.

A partir da análise dos dados históricos disponíveis, podemos concluir, em primeiro lugar, que a grande heterogeneidade etnolinguística da população de escravos africanos não criou uma situação favorável a interferências do substrato no processo de aquisição do português por essa primeira geração e na sua nativização, nas gerações seguintes. Por outro lado, a grande proporção de escravos em relação aos falantes nativos de português (destacando-se o expressivo contingente de escravos africanos) teria dificultado em muito o acesso aos modelos de português falado como língua materna, no processo de aquisição/nativização da língua entre a população escrava. Desse modo, pode-se pensar que a variedade de português transmitida para as gerações seguintes de escravos e seus descendentes, na região de Helvécia, tenha passado por profundas alterações, semelhantes às que definem o que se entende hoje por *crioulização*. E foram os traços remanescentes dessas profundas alterações que chamaram a atenção de Carlota Ferreira, pesquisadora do *Atlas Prévio dos Falares Baianos*, no início da década de 1960.

2.2.3 Caracterização linguística feita no início da década de 1960

Infelizmente, as duas pesquisadoras do *Atlas Prévio dos Falares Baianos* (*APFB*) não dispunham sequer de um aparelho gravador, quando chegaram a Helvécia, na noite de 25 de fevereiro de 1961. Após aplicarem os questionários no último ponto definido no *APFB*, a localidade de Ibiranhém, rumaram para uma localidade próxima, mas de difícil acesso, da qual tinham ouvido haver uma gente com uma fala "diferente". Chegando a Helvécia, confirmaram essas primeiras informações com os moradores mais jovens da vila, que "diziam que naquela cidade havia muita gente que falava diferente, 'engraçado', principalmente os mais velhos, e acrescentavam ainda que muitas vezes era difícil, para eles mesmos, filhos da terra, entenderem" (FERREIRA, 1984, p.22-23). Foi assim que, no dia seguinte, com lápis e papel na mão, as duas pesquisadoras do *APFB* buscaram registrar, junto a esses membros mais velhos da comunidade, algumas características de "um falar crioulo que deve ter sido geral, já que em 1961 dele subsistiam ainda vestígios" (FERREIRA, 1984, p.22).

Do pouco material que conseguiram recolher com dois informantes idosos (uma mulher de aproximadamente 75 anos e um homem de 80 anos), "foi na morfossintaxe, como é natural, que Helvécia nos forneceu maiores indícios de um possível crioulismo" (FERREIRA, 1984, p.28). E as seguintes estruturas foram registradas como evidências mais notáveis de um processo anterior de crioulização:

(i) uso variável do artigo definido

Ex.: "quando abri *janela*"

(ii) variação na concordância de gênero, tanto no interior do Sintagma Nominal quanto na relação com um termo predicativo:

Ex.: “io nõ póde rumá *o casa*”

“ela é muito *saído*”

(iii) simplificação da morfologia flexional do verbo:

- variação na flexão número-pessoal que atinge a primeira pessoa do singular:

Ex.: “io *sabe*”; “io *esqueceu*”

- uso da forma do presente pela forma do pretérito do indicativo:

Ex.: “io nõ *póde* rumá o casa” (‘eu não podia arrumar a casa’)

- uso da forma do infinitivo em contextos de formas finitas:

Ex.: “io *conhecê*” por ‘eu conheço’; “ele *morê*” por ‘ele morreu’; e quando io *andá* na Ponta de Areia, nõ tinha nada” (‘quando eu andava em Ponta de Areia, não havia nada lá’).

2.2.4 Conclusão: o diagnóstico do caso Helvécia e suas implicações para a história sociolinguística do Brasil

Os dados sócio-históricos relativos à composição da população de escravos da antiga Colônia Leopoldina configuram um cenário sociolinguístico muito propício ao surgimento de uma variedade linguística de base lexical portuguesa com uma estruturação gramatical muito afetada pelo contato entre línguas, num nível semelhante ao da formação das línguas crioulas típicas. A proporção entre escravos e brancos de dez para um é a que se estima para as situações prototípicas de crioulização. Nessa proporção, o acesso aos modelos da língua-alvo produzidos por falantes nativos é muito reduzido. O fato de muitos proprietários não serem falantes nativos do português certamente contribuiu, mas não parece ter sido decisivo nesse caso, porque é muito provável que os capatazes empregados para lidar com os escravos fossem em sua maioria brasileiros. Além disso, os escravos crioulos adquiridos de outros pontos da província, ou do país, viriam de zonas em que os senhores e capatazes eram falantes nativos do português. O mais importante é a elevadíssima proporção de escravos diante da população branca e a alta proporção, nesse segmento, de africanos.

Nesse contexto, ocorre uma violenta redução gramatical na variedade de português que é falada como segunda língua pelos escravos africanos. E, na medida em que essa variedade de português fornece o grosso dos dados linguísticos primários para aquisição da língua materna das crianças que nascem na comunidade, tem-se um cenário muito propício para o surgimento de uma variedade crioula de

português. No caso de Helvécia, essa nativização de um modelo defectivo de segunda língua parece ter sido o fator decisivo para a crioulização, suplantando uma pouco provável formação de um *pidgin* com base na transferência de estruturas gramaticais das línguas africanas, em face da grande heterogeneidade linguística do substrato reportada pelos documentos históricos.

Outro fator que parece ter sido decisivo no caso de Helvécia foi a permanência da grande maioria da população escrava na região após a abolição, associado ao seu grande isolamento até a primeira metade do século XX. Desse modo, teria sido de fato uma variedade crioula do português a língua que os informantes de Carlota Ferreira teriam adquirido como língua materna, já que eles nasceram entre os anos de 1880 e 1885, tendo, na altura do encontro, aproximadamente 80 anos, o homem, e 75 anos, a mulher. Mas, parece que o processo de descrioulização foi muito rápido, considerando-se que essa variedade crioula, no início da década de 1960, já era estranha aos membros mais novos da comunidade, que declararam ter, eles próprios, dificuldade para entender a fala "engraçada" dos mais velhos.

Infelizmente, as gravações da fala vernácula de moradores de Helvécia só vieram a ocorrer nos anos de 1987, quando Alan Baxter visitou a comunidade pela primeira vez, e de 1994, quando Alan Baxter e Dante Lucchesi recolheram a amostra que veio a integrar o *corpus* do português afro-brasileiro do **Projeto *Vertentes***, com o apoio do antropólogo Marcos Luciano Lopes Messeder e a inestimável ajuda de "seu" Quéte, um líder da comunidade, contando na época com aproximadamente 70 anos. Após 25 ou 30 anos, todos os antigos falantes do crioulo de Helvécia já haviam falecido, mas haviam deixado vestígios na fala dos que, com mais de 70 anos em 1994, tinham, no início da década de 1960, entre 35 e 55 anos. Esses indivíduos não falavam uma variedade crioula do português, mas exibiam processos de variação em sua fala, nos quais a forma padrão do português alterna com formas que teriam feito parte do repertório gramatical do antigo crioulo de Helvécia.

Algumas dessas formas, que foram registradas na fala crioula dos informantes idosos de Carlota Ferreira em 1961, ainda podiam ser detectadas, com uma frequência relativamente baixa, na fala dos entrevistados mais idosos de 1994. São elas: (i) a ausência de artigo ("*eu sô fia de lugá*", por '*eu sou filha do (deste) lugar*'); (ii) a ausência de concordância nominal de gênero ("*o meu sobrinha*"); (iii) a ausência de concordância verbal com a 1ª pessoa do singular ("*nunca mais eu sintiu*"). Outras formas que hoje são reconhecidas como características das línguas crioulas e que ainda ocorrem esporadicamente na fala dos informantes mais idosos da recolha de 1994 certamente eram as normais no antigo crioulo de Helvécia, sem que, contudo, tenham sido registradas por Carlota Ferreira. São elas:

(i) ausência de preposição em estruturas nominais: *folha mandioca* por *folha **de** mandioca*;

(ii) ausência do verbo copulativo: *esse aí neto de Casimiro* por *esse aí* ***é*** *neto de Casimiro*;

(iii) orações encaixadas sem complementizador: *ele disse a irmã dele veio do Rio*, por *ele disse* ***que*** *a irmã dele veio do Rio*;

(iv) negação verbal com sujeito marcado negativamente: *nenhum descarado num tá trabaiano não*;

(v) estruturas de duplo objeto com inversão na ordem dos complementos verbais e supressão da preposição de dativo: *Dá pessoá muita lembrança, minha fia!* por *Dá muitas lembranças ao pessoal, minha filha!*

(vi) uso de formas do presente para indicar ações e estados situados no passado: *meu pai é de cativeiro*, com o sentido de 'meu pai era escravo'.

E, assim como Carlota Ferreira anotou o uso de formas do infinitivo no contexto de formas finitas, o inverso também ocorria em 1994: *non sei vai lá* por *não sei ir lá*. Essas ocorrências que podem ser depreendias nas entrevistas dos informantes mais velhos de Helvécia servem como evidência empírica positiva para a hipótese de um processo de crioulização na formação histórica dessa comunidade de fala.

Se a crioulização em Helvécia, no século XIX, afigura-se como algo muito factível, a questão que se coloca, então, é saber o quão representativa ela é. O cenário de uma grande proporção de escravos, com uma presença significativa de africanos, em *plantations* relativamente isoladas do restante do país talvez não tenha sido muito comum na segunda metade do século XIX, mas, certamente, foi muito mais representativo no século XVIII e mais ainda no XVII. Com isso, reforça-se a hipotése, já aventada no capítulo anterior, de que processos localizados de pidginização e crioulização do português tenham sido mais frequentes nesses dois séculos, colocando-se como grande tarefa da pesquisa de linguística sócio-histórica mensurar, no conjunto da sociedade brasileira, ao longo da sua história, o quão representativo eram esses contextos propícios à crioulização.

Por outro lado, as informações provenientes de Helvécia apontam para um processo de descrioulização muito rápido. Algo em torno de uma ou duas gerações no máximo, uma vez que os moradores mais jovens de Helvécia, no início da década de 1960, já reconheciam a variedade linguística dos mais velhos como distinta da sua, fazendo referência, para Carlota Ferreira, de situações de ininteligibilidade entre as duas variedades linguísticas, que podem ser definidas aqui como o **crioulo português de Helvécia**, falado até as primeiras décadas do século, e o **português afro-brasileiro de Helvécia** que se generalizou em meados do século passado, como resultado da rápida descrioulização do primeiro.

A grande possibilidade de uma descrioulização rápida, nomeadamente no decorrer do século XX, vem a obnubilar a visão sobre a formação das comunidades

rurais afro-brasileiras, cujos registros de fala foram feitos só a partir da década de 1990, ou mesmo na primeira década deste século XXI. Tal é o caso das demais comunidades rurais afro-brasileiras isoladas aqui estudadas, para as quais não há elementos consistentes para se postular um processo pretérito de crioulização.

2.3 A comunidade de Cinzento

O termo *quilombo* (do quimbundo *kilombo*) foi usado primeiramente para designar fortificações dos jagas (designação dos povos que invadiram o Congo e Angola no final do século XVI). Por extensão, *quilombo* passou a nomear fortificações de negros fugidos do cativeiro, tendo, muitas vezes, como sinônimo a palavra *mocambo*, de origem controversa, mas provavelmente quimbundo (HOUAISS, 2001). Um quilombo era "um esconderijo de escravos fugidos" (MATTOSO, 2003, p.158), e tal prática seria uma das formas comuns de luta contra a escravidão. Tais espaços clandestinos constituíram formas grupais de resistência e de afirmação de uma comunidade unida não só pela cor, mas sobretudo pela luta contra formas as mais violentas de opressão e exploração. Normalmente, os quilombos não eram premeditados, nasciam espontaneamente, podendo reunir negros e crioulos, escravos ou homens livres (MATTOSO, 2003, p.158).

Pedro Tomás Pedreira (2001) queixa-se da falta de referência a quilombos baianos nos compêndios de História do Brasil, sendo os quilombos, em sua maioria, esquecidos pela História, a exemplo dos seguintes: o quilombo da Torre de Garcia d'Ávila, o quilombo de Jacuípe, o quilombo de Jaguaripe, o quilombo de Maragogipe, o quilombo de Muritiba, o quilombo de Cachoeira, o quilombo de Itaberaba (Orobó), o quilombo de Andaraí, o quilombo de Tupim (atual "Boa Vista do Tupim"), o quilombo de Xique-Xique e o quilombo do Buraco do Tatu (Cabula e arredores de Itapuã) (PEDREIRA, 2001, p.233). O conhecimento de tais quilombos advém de documentos escritos, localizados e datados, dando conta de sua constituição e do pedido de providências para que fossem destruídos. Pereira reconhece, também, a existência de outros quilombos, por via indireta, e supõe a presença de outros que não foram, ainda, listados.

Estima-se que a vinda dos primeiros negros para Cinzento tenha ocorrido na primeira metade do século XIX. Segundo relatos dos antigos moradores, os fundadores de Cinzento são da região da Chapada Diamantina, mais precisamente do antigo Arraial dos Crioulos, e o sobrenome *Pereira Nunes*, predominante em Cinzento, é o mesmo de um antigo proprietário de escravos da região de Rio de Contas. Ana Isidora, o membro mais velho da comunidade, com aproximadamente 107 anos de vida, afirma que os primeiros moradores vieram "currido", sugerindo

que a chegada se deu de forma clandestina, pois a vinda se deu “à meia-noite, terça-noite”.

A observação de Cinzento impõe de imediato a questão: por que seus fundadores escolheram um local tão acidentado para viver? Além disso, a falta de fontes naturais de água aumentava as dificuldades de sobrevivência. Associada a essa, surge uma outra interrogação: por que a comunidade esteve por tanto tempo isolada da cidade mais próxima? Curiosa, também, é a origem do nome que, por si, descreve o estado de pobreza dos seus primeiros moradores. Conta Ana Isidora que o nome provém de um boi gordo, “de tão gordo, parecia cinza”, que apareceu milagrosamente para alimentar os fundadores da comunidade. O que os levaria a abandonar o local onde viviam e se deslocar para uma terra com tantas adversidades como a de Cinzento? Certamente, fugiam da escravidão e buscavam no local a formação de uma irmandade com base primeiro na cor e depois no parentesco, já que a endogamia é uma prática comum até hoje.

Segundo o depoimento de Saviano Pereira Nunes, um dos membros da comunidade entrevistado na recolha da amostra de fala vernácula, foi Lourenço Pereira Nunes que desmatou as primeiras terras de Cinzento e delas tomou posse. Outro informante, Tercílio, afirmou-nos que os primeiros desbravadores foram o já citado Lourenço Pereira Nunes e Sérgio Pinheiro dos Santos. A valorização da terra não é apenas uma forma de sobrevivência particular, mas de todo o grupo, pois, da unidade territorial depende a unidade da coletividade. Esse aspecto explicaria a endogamia como forma de manutenção das terras na família, em caso de divisão.

Os mais velhos afirmam que seus pais contavam ter a obrigação de se ajoelharem “em frente dos donos”. Ana Isidora foi mais além e referiu que sua bisavó era “caboca do mato, e minha avó Maria foi pegada no mato”.

A comunidade de Cinzento é considerada atualmente uma comunidade afro-brasileira. Os cinzentenses são negros de pele bem escura. Todos são muito parecidos, em decorrência da endogamia em segundo grau. Graziele Ferreira (1999, p.70) descreve os cinzentenses da seguinte maneira: “fisicamente, são de estatura mediana, chegando até 1,70 m fortes, troncudos, de narinas alargadas, dentes largos e brancos, cabelos crespos e pés consideravelmente grandes”.

A situação dos afrodescendentes do Vale do Ribeira-SP descrita por Mary do Careno (1991, 2000) assemelha-se ao caso Cinzento. Em comum, pode-se ver o isolamento resultante da dificuldade de acesso à comunidade e da mudança advinda dos contatos com a cultura urbana. Além disso, como Careno (2000) acentua, a maior fonte de investigação encontra-se nos depoimentos e nas narrativas dos mais velhos como forma de resistência cultural. Em Cinzento, o caso do “boi cinzento” inclui-se na categoria de “narrativa fantástica”, nos moldes abordados por Careno (2000). Essa narrativa é a certidão de nascimento da comunidade, sendo contada e recontada por Ana Isidora, matriarca da comunidade.

2.4 Os arraiais de Rio de Contas

No Município de Rio de Contas-BA (13º 34' 44" de latitude sul e 41º 48' 41" de longitude oeste) foram recolhidas amostras de fala nas comunidades rurais afro-brasileiras isoladas de Barra e Bananal.

A ocupação da região de Rio de Contas se inicia na última década do século XVII, através do estabelecimento de uma rota de viagem entre Goiás e o norte de Minas e a cidade do Salvador, capital da então Província da Bahia. Com o intuito de estabelecer um "ponto de pouso", nessa rota de viagem, foi fundado um pequeno povoado com o sugestivo nome de *Creoulos*, situado em um planalto da Serra das Almas, na margem esquerda do Rio de Contas Pequeno, atual Rio Brumado.

Não tardou a descoberta de veios e cascalhos auríferos, não apenas no Rio de Contas, como também em seus afluentes e serras circunvizinhas. A fundação de Mato Grosso, três léguas acima do antigo povoado de Creoulos, subindo o Rio Brumado, a 1.450 metros de altitude, deu-se no bojo do grande afluxo de bandeirantes mineiros e paulistas para a região. Os jesuítas que acompanharam os bandeirantes ergueram no novo povoado uma igreja sob a invocação de Santo Antônio.

O desenvolvimento da mineração e o aumento da população do povoado foram de tal monta que, em 1718, foi criada a primeira freguesia do Alto Sertão Baiano — ou Sertão de Cima —, com denominação de Santo Antônio de Mato Grosso. Entretanto, em 1722, o Conselho Ultramarino decidiu criar, em função de uma carta dirigida ao Rei D. João V pelo Vice-Rei D. Vasco Fernandes César de Menezes, a Vila de Nossa Senhora do Livramento das Minas de Rio das Contas, doze quilômetros abaixo do antigo povoado de Creoulos, onde os jesuítas haviam erigido outra igreja, esta em devoção a Nossa Senhora do Livramento; e onde atualmente se situa a cidade de Livramento do Brumado. Uma nova reviravolta no povoamento da região aconteceria em 1745, quando uma Provisão Régia autorizou a mudança da sede da vila para o antigo povoado de Creoulos, que passou a se chamar Vila Nova de Nossa Senhora do Livramento das Minas do Rio das Contas, enquanto a antiga sede passou a ser conhecida por Vila Velha.

Essa mesma provisão também elevou a nova vila à categoria de freguesia, transferindo para aí a sede da Freguesia de Santo Antônio de Mato Grosso, com a denominação de Freguesia do Santíssimo Sacramento das Minas do Rio das Contas. Assim, enquanto Mato Grosso era deslocada para uma posição lateral mais isolada, para Rio de Contas iam afluindo todos os recursos e benefícios da atividade mineradora, o que se refletia no seu crescimento urbano. Lá foram construídos a Casa de Fundição, o Pelourinho e o edifício da Casa da Câmara e Cadeia Pública, este no início do século XIX, e até hoje conservado. Porém, com o progressivo esgotamento dos veios e cascalhos auríferos, o crescimento aos poucos vai se

estagnando. E, à medida que a atividade de mineração decrescia, diminuía também o nome do município. Em 1840, foi simplificado para Minas do Rio de Contas; e, em 1931, foram-se as minas, e o município passou a se chamar, simplesmente, Rio de Contas.

A origem das comunidades de Barra e Bananal, também chamadas Arraiais de Rio de Contas, deve ser coetânea ao início do povoamento da região pelos bandeirantes. Seus primeiros moradores foram possivelmente escravos foragidos que, em busca de um lugar seguro para sobreviver, seguiram o curso do Rio de Contas, estabelecendo-se nas cabeceiras do Rio Brumado, onde, desde então, praticam a agricultura de subsistência. Barra fica a uma distância de dois quilômetros de Bananal. Ainda os habitantes dos dois arraiais continuam a praticar a agricultura de subsistência e vivem em condições precárias, em termos de saneamento e educação. A endogamia é uma prática comum nos dois povoados.

A indústria do turismo instalada em torno da sede do Município de Rio de Contas veio quebrar o isolamento das duas comunidades. São frequentes as visitas de turistas e estudiosos, a fim de conhecerem os moradores dos dois povoados. Como ocorreu na comunidade de Cinzento, algumas das tradições de origem africana foram se perdendo, em função do contato com a cultura branca europeia, particularmente em função dos valores do catolicismo, que se tornou a religião predominante nas duas comunidades.

2.5 A comunidade de Sapé

A própria denominação da comunidade não é incontroversa. Os mais antigos chamam-na de *Sapé Grande*, enquanto os mais jovens de *Sapé Alto*, por influência do padre local, que não acha correto caracterizar *sapé* como *grande*, já que *sapé* é a designação de uma espécie de capim de folhas duras, muito conhecido por servir para cobrir choças e que é mal aceito pelo gado como forragem. Portanto, fica a polêmica linguística que se instaurou com a chegada do novo padre.

Distrito do Município de Valença, ao sul do Recôncavo Baiano, Sapé tem seu acesso por uma estrada de terra cujo percurso de 25 quilômetros se inicia em direção inversa à que leva à cidade de Valença, no entroncamento da BR-101.

Para os moradores da cidade, a comunidade formou-se algum tempo após a assinatura da abolição dos escravos. Segundo o depoimento de João Barreto, fazendeiro local e morador do distrito vizinho, Rapa Tição, em 1800 as terras que compreendiam Sapé, Rapa Tição e Tabuado pertenciam a um único dono, o Sr. Miguel Elia. Com a sua morte, as terras foram divididas entre seus dois filhos, que em seguida as venderam, por preferirem a vida urbana. Sobre as atividades econômicas da época, conta o Sr. João Barreto que se plantava largamente a

mandioca, além de haver áreas reservadas para o pasto, o que significa ter havido criação de gado, mas isso não afirma com certeza, pois, como ele mesmo diz, foram histórias que ouviu de seu avô, o primeiro a chegar à região após a família de Miguel Elia. Dizia seu avô que, ao tomar posse da fazenda, ainda nos últimos anos do século XIX, encontrou, nas terras, uma senzala e troncos onde os escravos eram castigados. Como registro da existência de escravos na região, o Sr. João Barreto ainda exibe alguns instrumentos dessa época, que guarda como relíquia. Mas as relíquias foram divididas com seus irmãos: as algemas estão em suas mãos, e as correntes, na fazenda de José Machado, seu cunhado.

Segundo os informantes, quando foram morar na região de Sapé, tiveram que desmatá-la. Não havia qualquer resquício de plantação, ou seja, ao chegarem, encontraram apenas sapé, o qual utilizaram para cobrir suas moradias. Então, o que levou essa pequena população a concentrar-se nessa região acidentada, sem recursos naturais e com o solo tão pobre?

Nessa região, mesmo após a abolição, muitos negros permaneceram nas fazendas, trabalhando sem nada receber. Os fazendeiros não pareciam gozar de uma boa situação econômica. Alguns anos depois, segundo o Sr. Nuna — um branco que foi morar próximo à região, em 1959, e buscou conhecer a formação da comunidade de Sapé —, os fazendeiros, para quitar suas dívidas com os ex-escravos, "doaram-lhes" pedaços de terra improdutivos, mais altos, sem trato e distantes da água. Esses ex-escravos ali permaneceram e casaram entre si, aumentando a população local, e, curiosamente, mantendo-se isolados. De acordo com o Sr. João Barreto, foram cinco negras de uma família de oito que deram origem à comunidade, entre elas, D. Isabel, mãe do Sr. Liordino, que diz ter espalhado pela região mais de 30 filhos.

Atualmente, a comunidade é formada por, aproximadamente, 100 habitantes, na região mais concentrada, e constituída, predominantemente, como em todas as sociedades mais recentes, por jovens e crianças. Apesar de haver energia elétrica há 10 anos, a maioria das famílias não tem o privilégio de utilizá-la, pois se restringe à rua principal. São poucos os moradores que possuem eletrodomésticos — geladeiras, por exemplo, há pouquíssimas, mas os aparelhos de TV são bastante encontrados, ligados à bateria de veículos em lugares não alcançados pela rede elétrica. Não há, na comunidade, saneamento e, nas casas, não há banheiros; quem os tem é considerado rico. Tanto homens, quanto mulheres e crianças trabalham na roça e ganham entre R$4,00 e R$10,00 a diária em época de plantio ou colheita. As famílias, em seus pedaços de terra, hoje bastante reduzidos em função da divisão entre os filhos, praticam agricultura de subsistência. Plantam feijão e mandioca, com a qual fazem farinha, cujo excedente vendem na BR-101. Possuem também alguns pés de cravo e cacau, de que tiram o mel para beber com cachaça ou para vender. Não há, na população local, a ambição de sair da comunidade

em busca de melhores condições de vida, como se pode constatar nas comunidades de Helvécia e Rio de Contas.

Quanto à religião, toda a comunidade é católica. Inquiridos sobre crenças africanas, os informantes afirmavam com orgulho não ter entre eles qualquer seguidor de cultos africanos. O isolamento dessa comunidade manteve algumas especificidades da fala local, tanto no nível fônico, como morfossintático.

2.6 Conclusão

A escassez de informações disponíveis acerca das comunidades de Rio de Contas, Cinzento e Sapé não possibilita afirmações mais conclusivas acerca da formação linguística dessas comunidades. Sendo as mais antigas, com as suas origens remontando a pelo menos a primeira metade do século XVIII, os arraiais de Rio de Contas seriam aqueles, para os quais se poderia pensar em processos de transmissão linguística mais intensos. Contudo, de acordo com argumentação desenvolvida no capítulo anterior, o contexto da mineração em que se formaram não seria o mais propício à crioulização. Quanto às comunidades de Cinzento e Sapé, por terem se formado muito provavelmente durante o século XIX, já se situariam em um período em que as condições para um contato entre línguas mais radical e profundo seriam menos frequentes. Portanto, diante dos registros disponíveis, a expectativa é a de que a comunidade de Helvécia seja aquela em que os efeitos de potenciais processos de mudança induzidos pelo contato entre línguas sejam mais notáveis.

Com efeito, somente na comunidade de Helvécia foram identificados processos de variação, em níveis estruturalmente significativos, que afetavam a concordância de gênero no interior do Sintagma Nominal e a concordância verbal junto à 1ª pessoa do singular; processos cujas origens estão, muito provavelmente, ligadas à transmissão linguística irregular desencadeada em situações de contato linguístico massivo e abrupto. Entretanto, essa diferenciação da comunidade de Helvécia nem sempre é nítida em processos de variação que afetam outros aspectos da gramática. De qualquer forma, só uma avaliação de conjunto dos aspectos da morfossintaxe analisados na segunda parte desse livro poderá levar a uma visão mais clara de como o contato entre línguas afetou a gramática das comunidades rurais afro-brasileiras isoladas. Para que essa avaliação seja eficaz, ela precisa mobilizar a teoria desenvolvida sobre o modo como as situações de contato entre línguas massivo, radical e abrupto afetam a estrutura linguística. Esse é o tema do próximo capítulo deste livro.

3 A transmissão linguística irregular

Dante Lucchesi
Alan Baxter

O conceito de **transmissão linguística irregular** é aqui tomado para designar amplamente os processos históricos de contato maciço entre povos falantes de línguas tipologicamente diferenciadas, entre os séculos XVI e XIX, em decorrência da ação do colonialismo europeu na África, Ásia, América e Oceania. Nas diversas situações de dominação que se constituíram nesse contexto histórico, a língua do grupo dominante, denominada **língua de superstrato** ou **língua-alvo**, se impõe, de modo que os falantes das outras línguas, em sua maioria adultos, são forçados a adquiri-la em condições bastante adversas de aprendizado, em função de sua sujeição e marginalização. As variedades de segunda língua que se formam nessas condições, mais ou menos defectivas consoante as especificidades de cada contexto histórico, acabam por fornecer os modelos para aquisição da língua materna para as novas gerações de falantes, na medida em que os grupos dominados vão abandonando as suas línguas nativas.

Tal processo de nativização da língua dominante ocorre de maneira irregular no sentido de que os dados linguísticos primários de que as crianças que nascem nessas situações dispõem para desenvolver a sua língua materna provêm praticamente de versões de segunda língua desenvolvidas entre os falantes adultos das outras línguas, que apresentam lacunas e reanálises em relação aos seus mecanismos gramaticais. Tal processo diferencia-se da situação de transmissão geracional normal das línguas humanas, em que as crianças dispõem de dados linguísticos mais completos fornecidos pela língua materna dos seus pais. Dessa forma, o processo de transmissão linguística irregular pode conduzir à formação de uma língua historicamente nova, denominada **língua crioula**, ou à simples formação de uma nova variedade histórica da língua de superstrato, que não deixa de apresentar processos de variação e mudança induzidos pelo contato entre línguas.

Este capítulo está estruturado da seguinte maneira. Na seção 3.1 é feita uma descrição do processo de formação de variedades linguísticas em situações de contato entre línguas massivo, radical e abrupto, visto como um processo que conjuga dois movimentos: a erosão e a reestruturação gramatical. Na seção

seguinte, são apresentados os condicionamentos sociais de tal processo. Já na seção 3.3 são detalhadas as características estruturais das línguas que se formam em tais situações de contato, contrastando as características das línguas *pidgins* com as das línguas crioulas. A seção 3.4 tem por objeto dois processos cruciais na reestruturação estrutural ocorrida em situações de transmissão linguística irregular: a gramaticalização e a transferência funcional. A natureza crioula de uma variedade linguística é discutida na seção 3.5, e a concepção da transmissão linguística irregular como um processo variável, o que se ajusta bem à compreensão da formação das variedades populares do português do Brasil em um contexto multilíngue, é o tema da seção que fecha este capítulo.

3.1 A formação de variedades linguísticas em situação de contato como um *continuum* de erosão e reestruturação gramatical

Quando uma população de adultos, falantes de línguas diferenciadas e mutuamente ininteligíveis, é forçada a adquirir uma segunda língua (L2) emergencialmente em função de relações comerciais e/ou de sujeição, a variedade dessa língua-alvo (LA) que se forma apresenta uma forte redução em sua estrutura gramatical, já que só os elementos essenciais, necessários ao preenchimento das funções comunicativas básicas, são mantidos. Essa redução na estrutura gramatical da língua deve-se:

(i) à dificuldade de acesso dos falantes das outras línguas aos modelos da LA, sobretudo nas situações em que o grupo dominante é numericamente muito inferior aos falantes das outras línguas;

(ii) ao fato de os falantes das outras línguas serem, em sua grande maioria, adultos, o que faz com que esses falantes não possam contar com os dispositivos inatos que atuam naturalmente no processo de aquisição da língua materna;

(iii) à ausência de uma ação normatizadora, ou seja, de uma norma ideal que oriente e restrinja o processo de aquisição da LA, já que esse processo tem seus objetivos circunscritos aos intentos comunicativos.

Dessa forma, o resultado desse contato linguístico massivo e abrupto é a coletivização de um conjuto reduzido de itens lexicais que cada falante das outras línguas manipula com base na sua competência linguística nativa, o **jargão**. O prolongamento da situação social de contato faz com que o jargão vá assumindo progressivamente novas funções na rede de interação linguística, ao tempo em

que se vai convertendo em modelo para a aquisição da língua materna dos descendentes dos falantes das outras línguas. Na medida em que o jargão, por suas limitações estruturais, é incapaz de atender às demandas decorrentes dessa expansão funcional, é preciso que haja também um incremento da sua estrutura gramatical, em um processo que se dá em torno da dialética entre a expansão funcional e a expansão gramatical, em função do desenvolvimento de uma rede de relações sociais e linguísticas que vão dar forma a uma nova comunidade de fala. A variedade linguística que se desenvolve em tal situação de contato é, portanto, fruto desse processo que conjuga dois movimentos:

(i) a erosão gramatical da LA, que, nos processos mais radicais de contato, pode levar à eliminação de todo o seu aparato gramatical no momento inicial de sua aquisição como segunda língua por falantes adultos; e

(ii) a recomposição gramatical dessa(s) variedade(s) defectiva(s) de segunda língua, que ocorre em função do estabelecimento da rede de relações sociais que vai dar ensejo à formação de uma nova comunidade de fala.

A compreensão de tal processo histórico de contato entre línguas radical e abrupto tem-se desenvolvido em torno das seguintes questões crucias:

(i) De que maneira a estrutura linguística é afetada, tanto no processo de erosão, quanto no processo de reestruturação gramatical?

(ii) Quem são os principais agentes desse processo? Os adultos, que recorrem ao repertório gramatical de sua(s) língua(s) nativa(s) para reestruturar gramaticalmente o código de comunicação emergencial? Ou as crianças, que criam uma língua nova a partir de um *input* muito precário em função dos dispositivos inatos da Gramática Universal que atuam no processo de desenvolvimento da língua materna?

(iii) Como esse processo é determinado por condicionamentos sociais, tais como: proporção numérica entre os membros do grupo dominante e dos grupos dominados, taxa de natalidade dos grupos dominados, grau de heterogeneidade linguística dos grupos dominados, nível de inserção/marginalização social dos falantes dos grupos dominados e seus descendentes, etc.?

Não obstante a importância dessa formalização esquemática para a compreensão do processo, as análises não podem perder de vista a sua grande variabilidade decorrente da combinação de todas as variáveis sociais envolvidas. Nesse sentido, até mesmo a ordem linear dos processos de erosão e reestruturação gramatical que figura em sua formalização não corresponde à realidade do processo em si, pois desenvolvimentos recentes na pesquisa sobre a **aquisição de segunda língua** (ASL) têm demonstrado que, em situações adversas, os falantes adultos tendem a desenvolver estágios aproximativos da LA, denominados **interlíngua**,

que não se caracterizam apenas por um processo de simplificação dessa língua, mas sobretudo pela criação ou desenvolvimento de um meio básico de comunicação verbal, no qual ocorrem transferências de dispositivos funcionais das línguas nativas dos adquirentes, bem como processos de reanálise dos mecanismos gramaticais da LA.

Isso não nega que a simplificação seja uma característica essencial da ALS em tais contextos (SIEGEL, 2008; KLEIN; PERDUE, 1997). Nesse sentido, deve-se destacar, em primeiro lugar, que mecanismos gramaticais mais abstratos e sem valor referencial[1] são normalmente eliminados, do mesmo modo os falantes adultos teriam maior dificuldade em adquirir mecanismos gramaticais da LA que estivessem ausentes em sua língua nativa, bem como fixar os parâmetros da LA distintos dos da sua língua (I. ROBERTS, 1999). No plano da Teoria da Gramática, a questão que se tem colocado é: *qual o nível de acesso dos falantes adultos aos dispositivos da GU nos processos de ASL?*[2] No caso de uma resposta negativa, esses falantes só teriam acesso aos dispositivos gramaticais e aos parâmetros fixados em sua língua nativa, de modo que a reestruturação gramatical assentaria nas transferências gramaticais da(s) língua(s) dos grupos dominados, denominadas **línguas do substrato**, para a variedade de L2 que se forma na situação de contato, bem como em processos de **reanálise** e **relexificação** (LUMSDEN, 1999a entre outros). Essa abordagem coloca os adultos como os agentes centrais do processo de reestruturação gramatical da variedade linguística que se forma na situação de contato, denominado processo de **expansão do *pidgin*** (MÜHLHÄUSLER, 1986; SIEGEL, 2008, p. 56-62), pois essa variedade de L2 que se desenvolve em situações de contato massivo e radical é tradicionalmente denominada ***pidgin***. Nesse sentido, deve-se destacar que o *pidgin*, que não é a língua materna da maioria dos seus falantes,[3] constitui uma variedade linguística qualitativamente distinta da LA, sendo ininteligível para os falantes monolíngues dessa língua. Não obstante a sua simplificação intrínseca, o *pidgin* pode alcançar uma estruturação gramatical semelhante à das línguas naturais, no ápice de seu processo de expansão, configurando o que se chama de ***pidgin* expandido** (MÜHLHÄUSLER, 1986).

Para que ocorra a pidginização é crucial que os falantes dos grupos dominados utilizem a(s) variedade(s) de L2, não apenas em seus contatos com os falantes do grupo dominante, mas sobretudo para se comunicarem entre si, no que Whinnom (1971) chamou de **hibridização terciária**. Nessa interação verbal

---

[1] Esses mecanismos correspondem, grosso modo, ao que Chomsky (1995) definiu como mecanismos não interpretáveis na interface semântica.

[2] Cf. I. Roberts (2007, p. 384-389).

[3] Cf. Muhlhäusler (1986, p. 5): "As línguas *pidgins* por definição não têm falantes nativos" (traduzido do original).

dos indivíduos dos grupos dominados, ocorre uma negociação em torno das soluções gramaticais empregadas para a formação de frases a partir do vocabulário básico da LA compartilhado por eles, e as estruturas gramaticais do *pidgin* vão se formando, na medida em que sejam consagradas pelo uso, dentro da rede de relações sociais que também se vão formando nessa comunidade emergente. Dessa forma, o *pidgin* vai se expandindo, na medida em que as regras gramaticais empregadas no seu uso são chanceladas dentro de uma rede de relações sociais historicamente determinada. No plano sócio-histórico, os *pidgins* tendem a desenvolver-se nos casos em que os grupos do substrato são mantidos em seu universo cultural de origem e conservam o uso de suas línguas nativas.

Nesse sentido, os *pidgins* emergem em situações de acentuado multilinguismo, como ocorreu em Papua-Nova Guiné, onde o *tok pisin* se desenvolveu nos últimos duzentos anos. Mais recentemente, variedades pidginizadas do inglês têm emergido em cidades africanas, particularmente na Nigéria.

A existência desses *pidgins* expandidos comprova que a reestruturação gramatical da variedade que se forma na situação de contato pode ocorrer independentemente da sua nativização, no que é tradicionalmente definido como **crioulização**.[4] Contudo, há toda uma corrente da crioulística para a qual a reestruturação gramatical depende crucialmente da crioulização (BICKERTON, 1981, 1984, 1988, 1999). Essa abordagem focaliza contextos sócio-históricos bem distintos, constituídos em torno de grandes empreendimentos agroex-portadores, denominados *plantations*. Nesses contextos, os falantes do substrato seriam retirados do seu universo social de origem e desarticulados cultural e linguisticamente, de modo que não seriam capazes de transmitir suas línguas nativas aos seus descendentes.[5]

Esse conjunto de características constituiria, segundo Bickerton (1999), a situação prototípica para a crioulização rápida, ocorrida no espaço de uma geração;

[4] Cf. o parâmetro adotado por De Camp (1971, p. 16) para definir língua *crioula*, como "a língua nativa da maioria de seus falantes" (traduzido do original), em oposição ao pidgin (cf. nota 2).

[5] Observações de campo realizadas pelos autores deste capítulo na Ilha de São Tomé, na África Ocidental, entre os descendentes de trabalhadores braçais trazidos do continente, no início do século XX, para trabalharem em grandes plantações de café e cacau, sugerem que a transmissão das línguas nativas do segmento do substrato se dá de maneira variada, podendo atingir indivíduos até da terceira ou quarta geração de nascidos na situação de contato (mesmo que apenas na forma de um conhecimento passivo). Isso não impede que, desde a primeira geração, os novos nascidos adquiram simultaneamente a LA através do modelo defectivo dos pais. Pelo contrário, o que se observa através das gerações é o fortalecimento da transmissão da língua socialmente hegemônica e um progressivo enfraquecimento da transmissão das línguas nativas do substrato, até o seu total desaparecimento. S. Roberts (2005) demonstra que, também no Havaí, a primeira geração nascida na ilha de trabalhadores de engenho contratados aprendeu as línguas ancestrais. Nesse caso, foi só com a geração seguinte, monolíngue, que as estruturas do crioulo foram fixadas.

uma situação sociolinguística que não possibilitaria o desenvolvimento gramatical do *pidgin*. Desse modo, as crianças que iam nascendo nas sociedades de *plantation* em formação dispunham, para o desenvolvimento de sua língua materna, de um *input* praticamente desprovido de estrutura gramatical. A teorização de Bickerton acerca do processo de crioulização assenta fundamentalemnte nas suas observações de campo feitas no Havaí.

O Havaí foi o cenário de um dos mais recentes processos de crioulização de que se tem notícia, em função do surto de desenvolvimento da cultura da cana-de-açúcar ocorrido no final do século XIX, provocando o afluxo de muitos trabalhadores braçais provenientes do Japão, da China, da Coréia, das Filipinas, de Portugal e de Porto Rico. Esses trabalhadores, em sua grande maioria já adultos, adquiriram o inglês de forma bastante precária, adaptando um parco vocabulário de palavras inglesas às estruturas de suas línguas nativas. Essa linguagem emergencial passou a ser o meio do intercurso verbal desses imigrantes, dando origem ao *pidgin inglês do Havaí*. Esse *pidgin*, com uma estrutura gramatical incipiente, diferia bastante da língua materna dos descendentes desses imigrantes que, não era o *pidgin*, nem o inglês, mas uma língua que, embora exibindo um léxico majoritariamente inglês, possuía uma gramática qualitativamente distinta da gramática da língua inglesa, com sistemas originais de artigos, de marcadores de tempo, modo e aspecto e de orações relativas, entre outras estruturas que Bickerton afirmou estarem ausentes no *pidgin*.[6] Essa gramática nova, a do *crioulo inglês do Havaí*, não era também uma mera colagem de estruturas gramaticais do chinês, do japonês ou do português, ou seja, das outras línguas envolvidas na situação de contato. Portanto, a gramática que teria emergido no processo de crioulização só poderia ser o resultado dos mecanismos atuantes no processo de aquisição da língua materna, que Bickerton (1981, 1984, 1988) denominou **Bioprograma da Linguagem**.

Em linhas gerais, essas seriam as formalizações existentes acerca da situação arquetípica dos processos consensualmente definidos como de **pidginização** e **crioulização**. Entretanto, os processos históricos de contato entre línguas são altamente variáveis em termos de seus parâmetros sociodemográficos e etno-linguísticos. Desse modo, pode-se pensar que, durante o período de expansão funcional/gramatical da nova variedade linguística, seja na pidginização, seja na crioulização — e mesmo depois, no chamado *continuum post*-crioulo (WINFORD, 2003) —, os processos de incorporação de dispositivos gramaticais das outras línguas e de reestruturação original da gramática sofrem uma concorrência dos

[6] Para Bickerton (1999, p. 70, nota 21), "em todo o período do estágio de jargão do *pidgin* no Havaí (aproximadamente, 1880-1920), não se desenvolveu, entre os adultos, qualquer representação de flexão, qualquer encaixamento de oração, qualquer sintagma complexo [...] ou qualquer outro índice de que eles tinham acesso a uma estrutura hierarquizada" (traduzido do original).

modelos da língua do superstrato que é mais ou menos intensa e determinante, conforme cada caso histórico particular (SIEGEL, 2008, p. 53-56). Daí a importância de se pensar o contato entre línguas nos termos mais amplos de um processo de transmissão linguística irregular, e não nos termos estritos das situações típicas de pidginização e/ou crioulização.

Nas situações típicas de pidginização/crioulização, o difícil acesso dos falantes das outras línguas aos modelos da LA mantém-se, geralmente, durante todo o período de formação da nova entidade linguística, que pode se tornar a LA para os novos falantes que ingressam na comunidade.[7] Portanto, a expansão gramatical dessa variedade linguística que se forma na situação de contato decorre principalmente dos processos de reestruturação original da gramática e da transferência de estruturas provenientes das línguas do substrato.

Porém, se o acesso dos falantes das outras línguas (e, principalmente, dos seus descendentes) aos modelos da LA aumenta com a continuidade da situação social que originou o contato, vai-se estabelecendo uma espécie de ideal normativo, na medida em que as formas da LA gozam de um maior prestígio social. Em tais contextos, os modelos da língua do grupo dominate tendem a suplantar os processos de transferência de estruturas das outras línguas e/ou de reestruturação original da gramática. O resultado, então, pode não ser a formação de um sistema linguístico muito distinto da LA, mas uma nova variedade histórica dessa língua que não deixa de apresentar certas características decorrentes do processo de transmissão linguística irregular que ocorreu com a sua socialização/nativização entre os falantes das outras línguas e seus descendentes. De qualquer forma, assim como na definição dos processos de pidginização/crioulização, a definição dos processos de transmissão linguística irregular deve conjugar os fatores linguísticos e psicolinguísticos com os fatores de natureza sócio-histórica.

3.2 Parâmetros sócio-históricos

No plano sócio-histórico, a situação que é considerada típica para a emergência de línguas *pidgins* e crioulas é aquela do deslocamento forçado de populações para um contexto geográfico e cultural distinto do seu contexto original. Tais processos ocorreram, de forma mais significativa e dramática, na história ocidental, entre os séculos XVI e XIX, durante o período da expansão colonial européia sobre os demais continentes, particularmente o continente americano,

[7] Deve-se ter em mente que as sociedades de *plantation* eram alimentadas no continente americano por um fluxo contínuo de escravos africanos trazidos pelo tráfico negreiro (cf. capítulo 1 deste livro).

denominado, então, o Novo Mundo. Nesse continente, os colonialistas europeus necessitavam de mão de obra para implementar os seus empreendimentos agroexportadores e de mineração. Diante da resistência cultural dos ameríndios ao trabalho forçado, os colonialistas encontraram na escravidão africana um meio mais eficaz — e muito lucrativo — de atender à enorme demanda de mão de obra do empreendimento colonialista.

A óbvia falta de controle sobre o tráfico negreiro não permite definir com precisão o número de indivíduos transportados — as estimativas variam entre 9 e 15 milhões quanto ao número de africanos trazidos para a América entre os séculos XVI e XIX. Todo esse contingente era proveniente de regiões com uma diversidade étnica e linguística muito grande; logo, era comum a reunião numa mesma localidade americana — uma mina, ou uma *plantation* — de africanos que não podiam se comunicar entre si usando as suas línguas nativas. Desse modo, o recurso a um código emergencial de comunicação verbal era, em muitos casos, uma necessidade não só para a interação entre os senhores e capatazes com os escravos, mas para a própria comunicação entre os escravos.

Nesses contextos, a proporção relativa entre a população de escravos e a população dominante desempenha um papel decisivo. Nas situações em que a proporção de escravos e seus descendentes crioulos fosse muito alta (em muitos casos, esses segmentos chegavam a constituir mais de 90% do total da população envolvida na situação de contato), o acesso aos modelos da língua dos segmentos dominantes seria bastante reduzido, favorecendo a ocorrência de processos de transferência de estruturas das outras línguas e de processos de reestruturação original da gramática na socialização/nativização da LA. Já nos casos em que a proporção de falantes da LA era maior, esses processos seriam mais raros e marginais.

Outro fator de grande importância para o resultado linguístico final da situação de contato seria a homogeneidade ou a heterogeneidade linguística da população escrava. Nos contextos de maior homogeneidade, pode-se imaginar que haveria uma maior transferência de estruturas das línguas dos segmentos dominados para a nova variedade linguística surgida na situação de contato. Já nos casos de maior heterogeneidade, o recurso à língua do grupo dominante tenderia a recobrir muito rapidamente uma gama bem mais ampla de situações de interação verbal, do que a simples interação imperativa entre senhores e capatazes, de um lado, e escravos, do outro. De qualquer modo, as línguas nativas tenderiam a ser progressivamente abandonadas, à medida que a população de escravos adquiria a língua do grupo dominante, mais viável no novo contexto social.

Nesse sentido, a continuidade da importação de novos escravos e a taxa de natalidade na população dominada durante o período de contato também seriam fatores externos a serem considerados. A importação de novos escravos implica a

reintrodução de variedades mais defectivas de L2 entre a população escrava, já que esses escravos teriam menos proficiência na LA do que os escravos antigos. Já a taxa de natalidade entre a população escrava teria uma implicação direta para a nativização da L2 emergente, o que pode ser decisivo no processo de transmissão linguística irregular.[8] Portanto, as línguas crioulas, embora manifestem semelhanças estruturais e funcionais, podem também manifestar consideráveis diferenças; consoante a maneira como esses diversos fatores externos se combinaram em cada situação histórica particular.

A noção da crioulização variável foi proposta de maneira explícita por Baker (1982) e Bickerton (1984), que afirmaram que a crioulização a partir de situações de contato com maior acesso aos modelos da LA deram origem a crioulos gramaticalmente mais próximos da LA. Baker baseou-se em estudos dos crioulos das Ilhas Maurício e de Reunião. Na Ilha de Reunião, Baker atribuiu a presença no crioulo de elementos de morfologia flexional do verbo ao fato de a proporção de escravos para europeus ter sido relativamente baixa durante as primeiras décadas da colônia. Essa situação contrasta com o que aconteceu na Ilha Maurício, onde, depois de um breve período em que os colonos prevaleceram em número, a população escrava aumentou rapidamente, até atingir uma ampla maioria. Nesse caso, desenvolveu-se um sistema verbal muito diferente daquele da LA, com uma estrutura morfológica analítica constituída de partículas temporais e aspectuais independentes.

Assim, a transmissão linguística irregular constitui um *continuum* de níveis diferenciados de socialização/nativização de uma L2 adquirida de forma mais ou menos imperfeita, em contextos sócio-históricos específicos. Já a crioulização típica se situa no extremo desse contínuo, numa situação em que o acesso à língua de superestrato foi extremamente reduzido, desencadeando um processo de reestruturação linguística independente, cujo resultado é a formação de uma nova língua qualitativamente distinta dessa língua. Porém, mesmo no setor crioulo do contínuo, a natureza dinâmica dos fatores externos que condicionam a transmissão linguística irregular definem resultados diversos, alguns mais radicais, outros menos radicais.

3.3 Aspectos linguísticos

A característica essencial da situação de contato linguístico massivo, abrupto e radical é a redução da gama das funções desempenhadas pela comunicação verbal (nos casos mais dramáticos, as funções sociais da língua ficam reduzidas às funções

[8] Essa questão será retomada na seção seguinte.

de comando e de comunicação básica) e uma perda de matéria gramatical da LA[9] que atinge sobretudo as estruturas mais abstratas, de menor funcionalidade comunicativa e de carga semântica mais tênue ou menos transparente,[10] comprometendo, em maior ou menor grau: as marcações morfológicas das categorias gramaticais de caso, número, gênero, pessoa, tempo, modo e aspecto, e as relações sintáticas de regência, concordância, coordenação e subordinação, bem como os processos de ligação anafórica e de movimento no âmbito da sentença. Desse modo, nesse processo de formação-aquisição da L2 que caracteriza a **pidginização**, os falantes das outras línguas produzem estruturas reduzidas e aproximações da língua de superstrato, realizando transferências funcionais a partir da gramática das suas línguas maternas (SIEGEL, 2008, p.26-104), e produzindo inovações originais por meio de gramaticalizações internas (WINFORD, 2003, p.278-286, 333-341). Entre as tendências estruturais que, no nível da morfossintaxe, são encontradas nas línguas *pidgins*, podem ser citadas as seguintes (MÜHLHÄUSLER, 1986, p.152-165; WINFORD, 2003, p.275-276):

(i) ausência de morfologia flexional;

(ii) tendência para a ordem SVO;

(iii) ordem invariável para as orações afirmativas, imperativas e interrogativas;

(iv) qualificadores sentenciais externos;

(v) sistema pronominal reduzido ao mínimo de formas;

(vi) uso reduzido de preposições;

(vii) ausência de regras de movimento, apagamento e relativização nos processos de derivação da sentença;

(viii) forma bimorfêmica analítica assumida pelas palavras interrogativas (constituintes WH).

Essas seriam tendências em línguas *pidgins* estáveis, ou seja, uma variedade linguística qualitativamente distinta da LA, que se desenvolveu gramatical e funcionalmente, sem se tornar língua nativa da maioria dos seus falantes, como aconteceu com o *tok pisin*, em Papua-Nova Guiné, e com outros *pidgins* da Oceania, nos últimos duzentos anos; assim como aconteceu com a expansão de alguns

---

9 Essa simplificação, de acordo com Muhlhäusler (1986), pode implicar simultaneamente uma redução do elenco de itens gramaticais e uma maximização funcional dos itens gramaticais supérstites.

10 Romaine (1988, p.28) chama a atenção para a substituição de formas sintéticas por formas analíticas, "de modo que formas complexas são decompostas em seus componentes". Ou seja, formas mais complexas e de difícil decodificação são substituídas por conjuntos de formas independentes mais simples e cujo significado é mais transparente. Nesse processo de otimização estrutural que caracteriza essa primeira fase da transmissão linguística irregular (MÜHLHÄUSLER, 1986), ocorre também "uma drástica redução da alomorfia" (ROMAINE, 1988, p.27).

pidgins em cidades africanas, mais recentemente. A possibilidade de esse processo ocorrer independentemente da nativização reforça a visão do caráter social do processo de reestruturação gramatical, num contraponto à visão daqueles que consideram que o processo de crioulização (a nativização da língua segunda surgida na situação de contato) é crucial para a recomposição da estrutura gramatical da nova variedade linguística.

Para aqueles que defendem que a crioulização é crucial nos processos de transmissão linguística irregular (BICKERTON, 1981, 1984, 1988, 1999), a aquisição da língua materna (AL1) que ocorre em situações de contato massivo é bastante diferente daquela que ocorre nos processos normais de transmissão de uma língua natural de uma geração a outra. No cenário "normal", uma geração de falantes fornece os dados linguísticos primários (DLP) para o processo de aquisição da nova geração a partir de seu desempenho em sua língua nativa, de modo que a língua materna dos pais é que fornece os modelos para a aquisição da língua materna dos filhos. Os parâmetros da gramática da criança são estabelecidos com base nesses dados, e a gramática adquirida se aproxima daquela da geração anterior, embora não seja idêntica a ela.[11]

Já no cenário de crioulização, os DLP para AL1 provêm de uma variedade de segunda língua que contém informações morfossintáticas altamente variáveis e defectivas. Nesse sentido, é fundamental o fato de o adulto que aprende uma L2 estar severamente limitado na sua capacidade de adquirir a morfologia flexional, sobretudo nas fases iniciais de aquisição da LA (DEGRAFF, 1999, p.517). A aquisição de uma L1 com base nesse gênero de DLP tem implicações muito sérias para a fixação dos parâmetros da gramática da L1 em desenvolvimento, e poderia dar conta do surgimento das estruturas morfossintáticas analíticas, muitas vezes observadas em línguas crioulas.[12] A incapacidade de fornecer dados da morfologia flexional para funcionar como estímulo-gatilho para a fixação de parâmetros teria implicações profundas para a fixação de parâmetros relativos às estruturas sintáticas associadas à morfologia flexional, tais como o parâmetro do sujeito nulo e as regras de movimento (I. ROBERTS, 1997).

---

[11] Seguimos a perspectiva da fixação de parâmetros delineada em Chomsky (1981) e bastante geral nos estudos aquisicionais. Segundo essa perspectiva, a criança adquire/desenvolve a sua gramática interna a partir de evidências/estímulos contidos nos DLP e das estruturas da Gramática Universal inata (GU). Os estímulos desencadeiam a fixação dos valores paramétricos especificados pela GU, definindo a gramática específica da língua adquirida naquela situação.

[12] Essas estruturas poderiam ser desenvolvidas a partir de evidências positivas nos modelos-estímulos disponíveis, ou poderiam partir de um dispositivo *default* da Gramática Universal, que entraria em jogo na ausência de evidências linguísticas claras no modelo-estímulo que indicassem parâmetros específicos (BICKERTON, 1999; LIGHTFOOT, 1999a; I. ROBERTS, 1999).

Bickerton (1981, 1984, 1988) apresenta as seguintes tendências estruturais das línguas crioulas, que seriam resultantes da expansão ocorrida com a nativização de um *pidgin* rudimentar:

(i) focalização do constituinte através do seu deslocamento para o início da oração;

(ii) sistema de artigos baseado na distinção entre os níveis de referencialidade 'específico/não específico', marcados, respectivamente pela presença/ausência do artigo;

(iii) sistema analítico, em que partículas independentes marcam as categorias verbais de tempo, modo e aspecto;

(iv) orações relativas com cópia pronominal;

(v) estruturas de negação dupla;

(vi) uso de um mesmo verbo para indicar 'posse' e 'existência';

(vii) ausência de cópula em predicados que envolvem adjetivos;

(viii) o uso de adjetivos que funcionam como verbos;

(ix) ordem invariável para as orações afirmativas, imperativas e interrogativas;

(x) forma bimorfêmica analítica assumida pelas palavras interrogativas (constituintes WH);

(xi) escassez de construções passivas.

Do cotejo dessas propriedades com as anteriormente apresentadas para as línguas *pidgins*, constata-se que há muito em comum entre esses dois tipos de língua. Diante disso e dos novos dados linguísticos do *pidgin* inglês do Havaí, S. Roberts (2005) e Siegel (2008) chegaram à conclusão de que vários traços da lista de Bickerton seriam o resultado de desenvolvimentos no *pidgin* prévio, em resposta a uma expansão das suas funções sociais como meio de comunicação interétnica. São achados que encontram paralelos nos trabalhos de Sankoff (1984, 1994) e Muhlhäusler (1986) sobre o *tok pisin*, que já tinham identificado expansões semelhantes às encontradas em línguas crioulas. Além disso, há ainda outros traços nessa lista que podem ser atribuídos a expansões que provêm de transferências funcionais a partir das línguas do substrato (SEIGEL, 2008, p.197-202), o que apontaria para uma fase em que ainda havia falantes das línguas ancestrais usuários do *pidgin*. Portanto, em determinados casos, pode ser difícil distinguir no processo de transmissão linguística irregular o que seria próprio à expansão do *pidgin* do que seria próprio à crioulização; tornando-se assim nebulosa a distinção que está na base da oposição entre aqueles que defendem que o processo de transmissão linguística irregular é regido por fatores socioculturais e aqueles que defendem que esse processo é fundamentalmente determinado por fatores psicobiológicos.

De qualquer forma, em ambos os casos, tanto na *expansão do pidgin*, quanto na *crioulização*, o que está essencialmente em jogo é a necessidade de recomposição das estruturas gramaticais perdidas na fase inicial do processo de transmissão linguística irregular; recomposição esta que, obviamente, será menor ou maior, consoante a intensidade do processo de erosão gramatical que se dá no início da situação de contato. Se as condições sociolinguísticas proporcionam um maior acesso aos modelos da língua de superstrato, menores serão a erosão e, consequentemente, a necessidade de recomposição da gramática; e quanto menor for esse acesso, maiores serão a erosão e o processo posterior de reestruturação gramatical. E o nível de acesso à língua de superstrato (condicionado por fatores socioculturais e demográficos) continua sendo determinante nas fases seguintes da expansão estrutural da nova variedade linguística (seja no plano da rede de relações sociais, seja no plano da aquisição da língua materna nas gerações seguintes), pois um menor acesso à LA abre espaço para a incorporação e fixação de elementos gramaticais do substrato.[13]

Entretanto, é evidente que a influência do substrato é uma variável que estaria condicionada pelo grau de homogeneidade linguística dos grupos dominados e/ou pelo estatuto social deses grupos. No plano histórico-cultural, as soluções do substrato teriam de vencer um grande obstáculo: o forte estigma que a herança cultural africana carregava nas novas sociedades coloniais, não sendo rara a repressão aberta a essas manifestações culturais e, principalmente, religiosas. O fato é que há crioulos com fortes e bem evidentes credenciais substratais, tais como o *palenquero*, crioulo de base lexical castelhana com substrato banto (quimbundo), falado numa comunidade perto de Cartagena, na Colômbia (SCHWEGLER, 1998), e o crioulo de base lexical francesa do Haiti, com substrato fon-ewe (LEFEBVRE, 1998). Em outros crioulos, a influência do substrato é menos evidente, por exemplo no crioulo cabo-verdiano.

Desse modo, tanto a *transferência funcional* como a *gramaticalização* desempenham papéis centrais, embora variáveis, na reestruturação da variedade linguística que se forma na situação de contato, seja essa reestruturação determinada por fatores internos no processo de nativização, seja ela determinada por fatores sociais.

[13] A influência da gramática da língua nativa dos falantes do substrato na formação da L2 tem sido tema de diversas investigações na área da Linguística Aplicada sobre a aquisição da L2, revelando que a **transferência** pode ter uma influência muito significativa (FRANCESCHINA, 2002; HAWKINS; CHAN, 1997), a depender da fase de desenvolvimento da L2 e a depender das circunstâncias sociais (SCHUMANN, 1978). O mesmo é válido para a área da crioulística, pois análises sobre o haitiano e o palenquero estão reacendendo a importância do substrato (JOURDAIN, 1985; KEESING, 1988, 1991; SIEGEL 2003, 2004a, 2008).

3.4 Mecanismos de expansão linguística envolvidos na transmissão irregular

A socialização e a nativização do código de emergência entre os segmentos dominados criam necessidades expressivas e comunicativas que demandam itens e estruturas para desempenhar papéis funcionais e expressar relações categoriais imanentes, desencadeando o processo de reestruturação gramatical da variedade linguística que se forma na situação de contato. Nos casos de pidginização e crioulização com o prolongado uso da variedade emergente, a necessidade de expansão estrutural frequentemente envolve um apoio na gramática da(s) língua(s) do substrato (L-sub). Nesse caso, pode haver incorporação de material morfológico da(s) L-sub,[14] ou pode haver uso de material lexical da língua de superstrato (L-sup) em funções morfológicas presentes (parcialmente ou plenamente) na(s) Lsub. Esse mecanismo é amplamente estudado e documentado por Siegel (2008, p.83-141, 148-202.), sob o título de **transferência funcional**.

A transferência funcional concorre com o processo de **gramaticalização** interna, que desempenha igualmente um papel de grande importância nos processos de pidginização/crioulização. A perda de matéria gramatical que se dá no momento do estabelecimento da situação de comunicação emergencial cria as condições bastante favoráveis à gramaticalização, num nível muito superior ao que se observa na história de variedades linguísticas que não passaram por situações de contato massivo, abrupto e radical. Desse modo, a recomposição gramatical da variedade linguística que se forma na situação de contato se dá em grande medida em função:

(i) da ampliação (ou maximização) das funções dos itens gramaticais da L-sup incorporados na L2 emergente / futura L1 da comunidade;[15]

(ii) da utilização de itens lexicais para expressar relações gramaticais.

Como exemplos do primeiro caso, encontram-se processos em que uma palavra gramatical assume, ao lado de sua função original, uma função ainda mais abstrata na estrutura da língua, como os demonstrativos *kel* (derivado do português *aquele*) e *se* (derivado do português *esse*), que podem eventualmente funcionar como artigo definido nos crioulos portugueses de Cabo Verde e São Tomé, respectivamente (LUCCHESI, 1993). Outros exemplos ocorrem quando uma mesma preposição provinda da L-sup passa a desempenhar funções que na L-sup são desempenhadas por duas ou mais preposições. Tal é o caso da preposição *ni* (< port. *em*), que

---

[14] Por exemplo, o palenquero incorporou o pluralizador prenominal quimbundo *ma-* (SCHWEGLER, 1998).

[15] Para uma exposição mais ampla sobre esta relação entre simplificação e maximização, veja-se Muhlhäusler (1986).

indica tanto localização ('lugar onde') como direção ('lugar para onde') no português dos tongas[16] e em variedades L2 do português de Angola (BAXTER, 2003) e Moçambique (GONÇALVES, 1996, 2004). Nesses casos, o efeito de transferências funcionais de línguas banto é óbvia, revelando a estreita relação entre os processos de transferência funcional e gramaticalização.

A conjugação desses dois processos está na base da formação dos sistemas analíticos de tempo, modo e aspecto (TMA) — uma das mais proeminentes características das línguas crioulas do Atlântico (BICKERTON, 1988; HOLM, 1988). As partículas de TMA que se fixam imediatamente antes do verbo como afixos gramaticais são derivadas de formas lexicais plenas. A partícula *te*, que no crioulo haitiano indica anterioridade, é derivada da forma *été* (ou *était*) do verbo francês *être* 'ser'. Já o aspecto durativo (ou progressivo) é indicado no crioulo cabo-verdiano pela partícula *ta*, derivada da forma *está* do verbo *estar*. Este aspecto era indicado no já extinto *Negerhollands Creole Dutch* através da partícula *lo*, derivada do verbo holandês *lopen* 'correr, andar, ir' (HOLM, 1988, p.156). Portanto, temos, nesses casos, formas lexicais plenas que perdem substância fônica e carga semântica, convertendo-se em afixos gramaticais.

Uma tendência frequente nas línguas crioulas do Atlântico, os tradicionalmente chamados *verbos seriais*, nada mais são do que fenômenos sintáticos constituídos fundamentalmente por uma transferência funcional que incorpora um verbo (um item lexical, portanto) que passa a funcionar como uma preposição ou complementizador; ou seja, como uma palavra gramatical/funcional.[17] Assim, no exemplo (1), abaixo, extraído do sranan, um crioulo falado no Suriname, o verbo inglês *give* 'dar', na forma reduzida *gi*, funciona como a preposição *para* (HOLM, 1988, p.184). Em krio, outro crioulo de base lexical inglesa, falado na Serra Leoa, na África, o verbo inglês *say* (*se*) 'dizer' funciona como complementizador junto a verbos do tipo *pensar*, *saber*, etc. — cf. exemplo (2). Já em sranan, o verbo usado é *talk* (*tak*) 'falar', cf. exemplo (3):[18]

(1) Kofi tjari den fisi ***gi*** mi

Kofi trouxe o peixe *para* mim.

[16] No português dos tongas, a preposição *ni* cobre as funções direcionais das preposições *em*, *a* e *de*.

[17] É importante ressalvar que, embora a serialização em outras línguas, por exemplo no chinês, seja um produto da gramaticalização interna, a presença de um verbo serial numa determinada língua não é necessariamente o produto de um processo de reestruturação puramente interno, porque, quando há substratos com serialização, o crioulo formado em contato com esse tipo de substrato costuma ter essa serialização também. Não está descartada, entretanto, a possibilidade de que a existência de verbos seriais em determinadas línguas crioulas seja uma consequência de um processo interno de gramaticalização.

[18] Exemplos extraídos de Holm (1988, p.185).

(2) a no ***se*** yu bizi.

Eu sei *que* você é ocupado.

(3) m sab ***tak*** a tru.

Eu sei *que* é verdade.

Porém, há casos mais radicais de transferência funcional, em que uma forma gramatical livre, no caso, uma preposição (*junto*), se converte em uma forma presa, o sufixo – *ntu*, marcador de caso no crioulo português do Sri Lanka. O mesmo acontece com o possessivo *sua*, que se transforma em um morfema de genitivo em estruturas mais basiletais como *mesa-su pano* 'pano de mesa', que se apresentam em variação com estruturas mais acroletais como *pano de mesa* (SMITH, 1977, p. 366-368). Nesses casos a influência do substrato tamil é inquestionável, como se pode ver nos exemplos (4) a (6) abaixo, retirados de Smith (1978, p.73), com a forma reduzida *pâ* da preposição *para*, que funciona como morfema de dativo, numa estrutura oracional muito mais próxima da encontrada na língua tamil, do substrato, do que a que se observa em português, a língua de superstrato:[19]

(4) port.: eu tinha dado o dinheiro para o João

(5) tamil: nan calli-yay jon-ukku kutu-tt iru-nt-an

eu dinheiro-ac. João-dat. pass.-dar-pass. aux.-foi-1ª pess.

(6) crioulo: eu dinhero jon-***pâ*** já-dá tinha

Outro caso de transferência funcional (esse a partir de substrato africano) é o uso da forma pronominal da terceira pessoa do plural como marcador de plural do SN, em muitos crioulos do Atlântico, como a forma *dem* (do inglês *them*) usada para indicar o plural dos nomes no crioulo inglês da Jamaica: *de pikini dem* 'as crianças' (KOEFOED, 1979, p. 39 apud ROMAINE, 1988, p.104), e no sranan: *den man* 'os homens'. Holm (1988, p.193) argumenta que esse plurarizador também funciona como marcador de definitude, como ocorre no iorubá, e "é tão freqüente nas línguas do grupo kwa e outras línguas do oeste africano e tão pouco usual nas línguas do mundo, que nenhum argumento sério foi proposto no sentido de atribuir essa estrutura crioula à influência de universais lingüísticos, em detrimento da influência do substrato".[20] Nesses casos, é interessante notar que a transferência funcional combina uma função presente na(s) língua(s) do substrato, que passa a ser desempenhada por um item lexical da língua do superstrato. No primeiro caso, preposições e pronomes do português se encaixam na morfologia nominal

---

[19] Os fatos históricos também sustentam essa influência do substrato, na medida em que a influência direta do português se extinguiu em 1658, muito cedo no processo de desenvolvimento do Crioulo (ROMAINE, 1988, p. 40).

[20] Traduzido do original.

de caso das línguas dravidianas; no segundo caso, um pronome do inglês se ajusta à estrutura de marcação de número no SN das línguas do Oeste africano.[21]

Pode-se pensar, então, em dois níveis de estruturação da variedade linguística que emerge no processo de transmissão linguística irregular. Num plano mais profundo, estaria a constituição das estruturas de base da gramática, em que concorrem os dispositivos da faculdade da linguagem com as estruturas das línguas de substrato e do superstrato que, consoante o contexto sócio-histórico, vão fornecer o *input* para o processo de expansão, e eventualmente, de nativização. Por outro lado, também na constituição dessas estruturas, ocorrerá a seleção dos itens lexicais que vão desempenhar suas funções e expressar os seus valores.

Nesse sentido, a partir de sua perspectiva inatista, Bickerton (1988, 1999) apresenta uma interessante análise, baseada na posição desenvolvida por Borer (1983) de que a fixação dos parâmetros sintáticos de uma língua (mesmo os mais básicos como a ordem de palavras na sentença) acontece no nível do léxico. Ou seja, as propriedades sintáticas da estruturação gramatical da língua se atualizam nas propriedades especificadas nas entradas dos itens lexicais. Assim, no processo de aquisição da língua materna pelas crianças, nas situações de transmissão linguística regular de uma geração a outra, "as propriedades dos itens lexicais que elas adquirem combinam com a sintaxe invariante, do que eventualmente resulta a aquisição de uma gramática preexistente".[22] Bickerton (1999, p.57), então, estabelece uma distinção entre o que acontece nesses casos e nos casos de transmissão linguística irregular (particularmente nos casos de crioulização):

> No caso normal, a criança de quatro ou cinco anos já terá adquirido uma ampla gama de itens gramaticais, suficiente para atender aos requerimentos (em termos de regência, anáfora etc.) definidos pela sintaxe inata. No caso crioulo, para a maioria desses requerimentos, a criança simplesmente não pode encontrar no *pidgin* os itens gramaticais apropriados. Portanto, os itens gramaticais terão de ser criados pelo recrutamento de itens lexicais, com o enfraquecimento do significado lexical destes últimos.[23]

Obviamente que a gramaticalização e a transferência funcional geram estruturas que não são privativas das línguas *pidgins* e crioulas. Assim, por exemplo, o processo de gramaticalização que faz com que um demonstrativo passe a desempenhar as funções de um artigo, que ocorreu na crioulização do português em África (como referido acima), também ocorreu na passagem do latim ao

[21] Essa concepção de transferência funcional é bastante desenvolvida por Lumsden (1999a), com nome de **relexificação**.

[22] Traduzido do original.

[23] Traduzido do original.

português, cujo artigo definido é derivado da forma do acusativo de um demonstrativo latino. Da mesma forma, a presença de verbos que passam a funcionar como preposições, constituindo as estruturas denominadas de verbos seriais nas línguas crioulas, também é documentada na história do chinês mandarim (LI, 1975). A diferença residiria, então, na quantidade e na intensidade das mudanças ocorridas num lapso tão curto de tempo no processo de transmissão linguística irregular *vis-à-vis* ao ritmo dessas mudanças no desenvolvimento, digamos, normal da língua. Em primeiro lugar, as inovações que surgem quotidianamente na interação verbal dos falantes de uma língua fora da situação de contato têm que vencer a força das formas e estruturas fixadas na tradição do uso linguístico, para que uma mudança possa ocorrer.[24] Com isso, mudanças desse tipo são muito limitadas, se comparadas ao que acontece no contato abrupto e radical entre línguas que constitui a gênese do processo de transmissão linguística irregular. A erosão da estrutura gramatical que costuma acontecer nesses casos produz uma enorme pressão estrutural, que abrirá caminho para uma legião de processos de gramaticalização e transferência funcional. Esse conjunto de mudanças é de tal monta que, nos casos mais extremos, determinará, no curto período de uma ou duas gerações, uma reestruturação original da gramática tão ampla e profunda que terá por consequência o surgimento de uma nova língua histórica.

Esse aspecto é crucial para que se possa entender o caráter e a natureza das línguas crioulas frente às chamadas línguas naturais. As reflexões mais aceitas no âmbito da crioulística apontam no sentido de que não se pode falar em propriedades e estruturas privativas das línguas crioulas; de modo que o caráter crioulo ou crioulizante de uma determinada variedade linguística não se define por suas propriedades estruturais imanentes, mas pelo processo histórico de contato linguístico que levou a sua formação *vis-à-vis* ao desenvolvimento histórico interno da língua que forneceu a sua base lexical. Portanto, podemos afirmar, com I. Roberts (1997, p.11), que “não há nada de especial nos crioulos, além da sua história”.[25]

3.5 A natureza crioulizante de uma variedade linguística

A natureza crioula de uma determinada língua é sempre definida em termos relativos e nunca de forma absoluta ou universal, pois todas as características estruturais e mudanças que ocorrem nas línguas crioulas podem ocorrer em outras

[24] Nesse sentido, os períodos de instabilidade e de enfraquecimento das instituições reguladoras do comportamento e das relações sociais tendem a favorecer a ocorrência de um número maior de mudanças linguísticas.

[25] Traduzido do original.

línguas naturais, não crioulas. E nem poderia ser diferente, na medida em que as línguas crioulas também são línguas humanas naturais. Uma língua crioula já formada desempenha virtualmente as mesmas funções comunicativas e sociais que qualquer outra língua natural e é transmitida normalmente de geração para geração, sendo adquirida pelas crianças com base na faculdade humana da linguagem como qualquer outra língua. Portanto, as suas características estruturais e funcionais fazem parte do leque de soluções geneticamente codificadas para todas as línguas humanas. Assim sendo, o que define primariamente uma língua crioula é o processo acelerado de reestruturação da gramática em sua formação, muitas vezes, abrupta. Portanto, o grau de crioulização de uma língua pode ser mensurado relativamente com base no grau de reestruturação gramatical divergente dos modelos disponíveis na língua lexificadora. Quanto mais divergentes forem as estruturas constituídas no processo de recomposição/reestruturação gramatical *vis-à-vis* aos modelos da LA, maior será o grau de pidginização e/ou crioulização.

Com base na Teoria dos Princípios e Parâmetros, I. Roberts (1997) defende que as línguas crioulas tendem a apresentar parâmetros de valores não marcados, mesmo quando esses parâmetros exibem valores marcados em sua língua lexificadora. I. Roberts (1997, p.11) parte do

> [...] princípio de que todo parâmetro depende dos valores dos traços abstratos dos núcleos funcionais: a variação paramétrica decorre do fato de um dado núcleo funcional H poder desencadear movimento de vários tipos. Se H desencadeia movimento, cria-se uma representação relativamente complexa. O valor desencadeador de movimento de um parâmetro (o valor 'forte' na terminologia de Chomsky) é então o valor marcado.[26]

Do ponto de vista da aquisição, "vê-se que só os valores marcados precisam ser expressos; no caso da não-expressão de um parâmetro, o *default* será o valor não marcado".[27] Desse modo, os crioulos apresentariam valores não marcados de parâmetros sintáticos em função das circunstâncias em que ocorre a sua gênese. Tendo como ponto de partida para a sua formação um *pidgin*, ou mesmo um mero leque de interlínguas (soluções individuais), o crioulo tende a se constituir com parâmetros fracos, em função da perda de morfologia que se dá na fase inicial do processo de transmissão linguística irregular. O estímulo-gatilho para a aquisição das gerações iniciais de falantes do crioulo pode ser "morfológica e sintaticamente defectivo em pontos cruciais" em relação a "certas propriedades requeridas para o desencadeamento de traços fortes" (I. ROBERTS, 1997, p. 12-13). Assim, muitas propriedades típicas das línguas crioulas

---

[26] Traduzido do original.

[27] Traduzido do original.

— tais como: a ausência de movimento do verbo, de sujeitos nulos referenciais e de clíticos complementos, bem como a ordem SVO e o sistema de partículas de tempo, modo e aspecto — podem ser analisados por essa ótica.

O processo é particularmente notável nos casos dos crioulos que apresentam um valor não marcado para um parâmetro que possui um valor marcado na língua lexificadora. O crioulo do Haiti, por exemplo, não apresenta o movimento do verbo para INFL presente na sua língua lexificadora, o francês (DEGRAFF, 1994; DEGRAFF; DEJEAN, 1994 *apud* I. ROBERTS, 1997). A falta de movimento está relacionada com a perda de morfologia flexional do verbo ocorrida na formação do crioulo, constituindo assim o valor não marcado do parâmetro (I. ROBERTS, 1997, p. 14). A perda de clíticos, também exemplificada na relação histórica entre o crioulo do Haiti e o francês, pode igualmente ser vista sob essa perspectiva. A existência de clíticos complementos constitui um parâmetro marcado, pois implica necessariamente uma operação de movimento aparente, o que não acontece na ausência dessas partículas. Pode-se, então, esperar que os crioulos não possuam esse tipo de clítico (I. ROBERTS, 1997, p. 19-20).

A ordem SVO, bastante geral entre as línguas crioulas (BICKERTON, 1981, 1988; MÜHLHÄUSLER, 1986), pode ser vista também como um valor não marcado do parâmetro, já que:

> [...] a ordem VO é a única ordem subjacente disponível e a ordem OV é derivada por movimento de DP para Spec, AgrOP para checagem de traço; então, OV resulta da presença de traços N fortes em AgrO, e o traço forte de AgrO acarreta representações mais complexas para a aquisição, na medida em que provoca mais movimento aparente; sendo, portanto, marcada (I. ROBERTS, 1997, p. 14-15).[28]

A ordem SVO do *Berbice Dutch*, derivado do holandês e do ijo (duas línguas de ordem OV), exemplifica essa passagem de um valor marcado para um valor não marcado no processo de crioulização. O parâmetro do sujeito nulo pode ser considerado como marcado, já que depende de uma morfologia flexional forte do verbo para permitir o reconhecimento do conteúdo referencial de *pro* (I. ROBERTS, 1997, p.16). Como ocorre a perda de morfologia flexional no processo de transmissão linguística irregular, vários crioulos, como o papiamento (de base lexical espanhola), não apresentam sujeitos nulos, enquanto suas línguas lexificadoras exibem essa propriedade. Porém, essas propriedades que caracterizam as línguas crioulas não são exclusivas dessas línguas, pois:

> [...] línguas não crioulas podem ter constelações similares de valores não marcados: o inglês, por exemplo, não tem movimento de verbo para

[28] Traduzido do original.

> os domínios funcionais mais altos, apresenta ordem SVO, não possui sujeito nulo, nem clíticos complementos, e ainda apresenta um sistema de marcadores pré-verbais de tempo, modo e aspecto bastante similar ao encontrado nos crioulos.[29]

Assim sendo, não se pode atribuir um caráter crioulo a uma mudança ou a uma estrutura linguística sem o recurso à relação histórica particular subjacente, determinada por contexto social de contato linguístico abrupto e massivo.

De qualquer forma, as estruturas que derivam de mudanças ocorridas em situações de contato linguístico resultam da não aquisição/incorporação de morfologia e elementos gramaticais da língua de superstrato, que ocorre na fase inicial do processo de transmissão linguística irregular, combinada com a reestruturação gramatical da nova variedade linguística. Portanto, os processos que caracterizam a formação de uma variedade linguística numa situação de contato propícia à transmissão linguística irregular são fundamentalmente:

(i) não aquisição/incorporação, ou variação no uso, de morfologia flexional e palavras gramaticais;

(ii) gramaticalização de itens lexicais para preencher as lacunas na estrutura linguística;

(iii) transferência funcional que acomoda itens lexicais do superstrato em funções gramaticais da(s) língua(s) do substrato; e

(iv) alteração dos valores dos parâmetros sintáticos em função de valores não marcados.

A definição dessas características fundamentais do processo de transmissão linguística irregular é crucial para poder avaliar com mais precisão as reais consequências linguísticas dos processos históricos de contato entre línguas que não resultaram na formação de um crioulo típico, mas que nem por isso deixaram de interferir decisivamente na formação das variedades da LA que sugiram nesse tipo de situação.

3.6 A transmissão linguística irregular como um processo variável

Os processos de transmissão linguística irregular que conduziram à formação de *pidgins* e crioulos ocorrem normalmente em contextos sócio-históricos que combinam os seguintes parâmetros:

---

[29] Traduzido do original.

(i) o acesso mais restrito à LA na fase inicial do contato entre línguas; bem como

(ii) a pouca influência da LA nas fases seguintes de reestruturação da gramática.

A consequência de (i) é uma maior erosão na estrutura morfológica –– o que favorece também uma maior alteração nos parâmetros sintáticos ––, e a consequência de (ii) é o desencadeamento de um processo mais autônomo de recomposição da gramática aparente com um número maior de processos de gramaticalização. Mas, mesmo entre as línguas crioulas, observam-se graus variados de reestruturação gramatical, com casos de reestruturação mais profunda, como, por exemplo, o crioulo francês do Haiti e o angolar — crioulo português da Ilha de São Tomé ––, e casos de reestruturação menos profunda, como, por exemplo, o crioulo francês da Ilha de Reunião e o crioulo cabo-verdiano. Isso se dá em função de uma maior ou menor intensidade dos parâmetros (i) e (ii), apresentados acima. De qualquer forma, mesmo nos processos de crioulização, as formas originais decorrentes dos processos de gramaticalização sempre enfrentam uma concorrência, maior ou menor, consoante as condições sócio-históricas, das formas da língua de superstrato. Mesmo após a consolidação da língua crioula, ainda se observa a influência das formas da língua de superstrato, que, em muitos casos, conduz à substituição das estruturas tipicamente crioulas pelas da língua de superstrato, num fenômeno conhecido como **descrioulização**, que se verifica no chamado ***continuum post*-crioulo**.

Quanto aos processos de transmissão linguística irregular que não resultaram na formação de *pidgins* ou crioulos estáveis, situação que mais se aproxima das que se observam no processo de constituição histórica da língua no Brasil (cf. seção 1.3 do primeiro capítulo deste livro), pode-se pensar nos seguintes parâmetros. Nos estágios iniciais do contato entre línguas, a erosão gramatical ocorre num grau menor e muito variável, do que resulta, não a eliminação, mas uma variação, maior ou menor, conforme o caso, no uso dos morfemas flexionais e dos elementos gramaticais; o que não impede necessariamente que ocorram opacidades nos estímulos-gatilhos que conduzam a alterações na marcação dos valores dos parâmetros sintáticos. Nos estágios seguintes, mantém-se o quadro de variação estruturada que caracteriza os padrões coletivos de uso linguístico, e, nos casos em que cabe uma recomposição da gramática, a tendência predominante seria a de uma reintrodução decorrente da influência da LA. Os processos originais de reestruturação gramatical e de transferência funcional seriam raros e periféricos.

Contudo, ao longo de todo o processo de transmissão linguística irregular, surgem, no plano heterogêneo da fala, os embriões de processos potenciais de gramaticalização. Na maioria dos casos, essas formas e estruturas sucumbem diante da influência da LA, mas elas podem também se manter em variação estruturada

com as variantes da LA por longos períodos. Por outro lado, no que se refere à alteração dos valores dos parâmetros sintáticos, pode-se pensar também numa situação de variação em que convivem, na comunidade de fala, duas ou mais gramáticas, em função da marcação diferenciada dos valores de seus parâmetros sintáticos. Portanto, os processos de transmissão linguística irregular mais leves, que aqui se postulam para explicar a história das variedades populares do português do Brasil e do português afro-brasileiro em particular (cf. capítulo 2 deste livro), compreenderiam as seguintes etapas:

(i) fase inicial: variação, mais ou menos intensa, no uso dos mecanismos gramaticais, sendo mais rara a eliminação desses mecanismos; opacidade relativa nos estímulos-gatilhos que possibilitam a aquisição de valores diferenciados nos parâmetros sintáticos;

(ii) fases seguintes: manutenção do quadro de variação, em que se observa a concorrência entre as formas gramaticais reintroduzidas a partir da influência da LA e potenciais processos de gramaticalização que emergem na heterogeneidade da fala; variação nas frequências de uso dos parâmetros sintáticos.

Portanto, para se determinar a influência dos processos de transmissão linguística irregular no muito vasto e complexo universo sociolinguístico brasileiro, é preciso levar em conta, por um lado, o item da estrutura linguística que está sendo focalizado, a sua função mais ou menos abstrata, sua transparência semântica e o valor mais ou menos marcado do parâmetro sintático em jogo; tendo, sempre, como referência os modelos transplantados da variedade linguística europeia; e, por outro lado, a comunidade de fala em que esse item é analisado, o que implica considerar o grau de contato linguístico por que essa comunidade passou, a sua exposição aos modelos linguísticos difundidos institucionalmente, etc.

Entretanto, há um aspecto básico do processo de transmissão linguística irregular que é crucial para o entendimento do desenvolvimento histórico do português do Brasil. O ponto de partida de todo processo de transmissão linguística irregular desencadeado pelo contato entre línguas é a perda de morfologia flexional na aquisição inicial da LA por parte dos falantes das outras línguas. Esta característica está na base da formação de todas as línguas *pidgins* e crioulas, e muitas das propriedades dessas línguas decorrem de mudanças desencadeadas por esse "fato inaugural". Muitas línguas crioulas não apresentam sistemas de partículas de tempo, modo e aspecto, nem possuem verbos seriais, mas todas as línguas *pidgins* e crioulas exibem algum nível de perda de morfologia flexional em relação à LA. Essa é uma das poucas assertivas absolutamente consensuais entre os crioulistas, como se pode inferir desta afirmação de Goodman (1971, p.253) entre tantas outras que se poderiam arrolar: "Um traço que é virtualmente universal

entre essas línguas geralmente classificadas como *pidgins* e crioulas é a drástica redução da complexidade morfológica e da irregularidade".[30]

Pode-se concluir, então, que *em todo e qualquer processo de transmissão linguística* irregular desencadeado pelo contato entre línguas ocorre, em maior ou menor grau, perda de *morfologia flexional*. Nas situações típicas de crioulização, essa perda tende a ser total. Já nos casos de uma transmissão linguística irregular mais leve (a exemplo do crioulo francês da Ilha de Reunião, o inglês de Singapura, o português de Maputo, em Moçambique, e o português dos tongas de São Tomé), não se configura um quadro de eliminação em níveis tão categóricos, mas um quadro de variação mais ou menos intenso conforme cada caso. Com base nessa fundamentação teórica, pode-se conceber o quadro de profunda variação na concordância nominal e verbal que se observa hoje nas variedades populares do português do Brasil como uma decorrência do processo de transmissão linguística irregular que marca a gênese dessas variedades linguísticas.

Assim, os fatores sociais e linguísticos se combinam de uma forma bastante sugestiva no sentido de que o contato entre línguas tenha desempenhado um papel relevante na história da língua portuguesa no Brasil. Do ponto de vista sócio-histórico, documenta-se uma presença representativa de falantes de outras línguas autóctones e, sobretudo, africanos que adquiriram o português já adultos, em situações bastante adversas; seguindo-se a socialização e a nativização dessa variedade segunda do português nesses segmentos. Por outro lado, observa-se ainda hoje um conjunto de processos de variação e mudança que se ajustam a um quadro de transmissão linguística irregular, sobretudo se focalizarmos as variedades populares do português brasileiro, particularmente as comunidades rurais afro-brasileiras isoladas, que constituem o universo de observação das análises contidas neste livro. Conforme argumentado no capítulo anterior, essas comunidades, em função de suas especificidades históricas, podem se constituir em um verdadeiro laboratório para se observar como o processo de transmissão linguística irregular afetou a gramática do português brasileiro. Para tanto, as análises que se apresentam na segunda parte deste livro buscam conjugar os padrões variáveis de uso linguístico com os processos de mudança que se operam na estrutura da gramática subjacente aos atos de fala, de modo a identificar como o processo de transmissão linguística irregular afetou a gramática dessas comunidades de fala em sua formação. Esse objetivo coloca mais um desafio para análise: integrar em uma abordagem variacionista os fundamentos teóricos mais gerais da teoria da gramática. Esse é o tema do próximo capítulo deste livro.

[30] Traduzido do original.

4 Teorias da estrutura e da mudança linguísticas e o contato entre línguas

Dante Lucchesi

Ilza Ribeiro

O estudo da mudança linguística a partir da variação observada sincronicamente em um estado de língua ganhou impulso na década de 60, com o surgimento do programa de pesquisa da **Sociolinguística Variacionista**, segundo o qual a análise da variação no seio de uma comunidade de fala poderia fornecer informações valiosas acerca de como uma determinada mudança se estaria implementando na estrutura linguística e social dessa comunidade. Além de possibilitar um diagnóstico preciso dos mecanismos atuantes nos processos de mudança em curso no presente, esse modelo postulava também poder fazer projeções para processos de mudança ocorridos no passado, com base no *princípio uniformitarista*, segundo o qual os fatores que regulam a variação e a mudança no presente seriam basicamente os mesmos que teriam atuado no passado.[1]

Paralelamente, uma concepção diferente da mudança linguística se desenvolveu no âmbito de um outro programa de pesquisa, a **Gramática Gerativa**. Nesse modelo, a mudança é regulada pela forma como o conhecimento linguístico se estrutura na mente dos falantes e é impulsionada crucialmente no processo de aquisição da língua materna (LIGHTFOOT, 1979, 1991, 1999b, entre outros). Desse modo, a mudança não seria regulada apenas pela estrutura gramatical de uma língua particular, mas também por estruturas inatas, comuns a todos os indivíduos da espécie humana, formalizadas através da ideia de uma Gramática Universal (CHOMSKY, 1957, 1965, 1981, 1986, 1995, entre outros).

---

[1] A reflexão sobre a Sociolinguística Variacionista baseia-se nos textos balizadores do programa de pesquisa, a saber: Weinreich, Labov e Herzog, 1968 (doravante WLH); Labov, 1966, 1972a, 1972b, 1982 e 1994. Os temas aqui abordados são recorrentemente tratados nesses textos. Assim, só será feita a remissão a um desses textos em particular, caso essa remissão venha a acrescentar algo ao raciocínio desenvolvido. Uma exposição mais abrangente e detalhada do desenvolvimento teórico da Sociolinguística Variacionista pode ser encontrada em Lucchesi (2004c).

Não obstante as profundas divergências que os separam, esses dois modelos têm produzido contribuições decisivas e complementares para o entendimento da questão da mudança, tanto no plano dos princípios teóricos fundamentais, quanto no plano da metodologia e dos critérios de verificação empírica. A contribuição desses dois modelos para a compreensão da questão da mudança constitui o objeto deste capítulo, que busca fundamentar uma posição teórica que permita integrar elementos da teoria da gramática e a análise dos padrões coletivos de uso linguístico de uma forma minimamente consistente. Por outro lado, considerando a questão das mudanças derivadas do contato entre línguas e da convivência de variedades linguísticas distintas dentro de uma mesma comunidade de fala, central dentro da temática deste livro, buscar-se-á uma formalização analítica dos processos em que gramáticas concorrentes afetam o desempenho linguístico dos falantes, tratadas na literatura sob o rótulo de **alternância de códigos** (ing. *code-switching*), como um processo encaixado na estruturação gramatical do conhecimento linguístico dos falantes, através do que se definiu como **princípio da coesão estrutural**.

4.1 A teoria da gramática e o estudo dos padrões coletivos de comportamento linguístico

A definição de **comunidade de fala** como objeto de estudo da Sociolinguística Variacionista está baseada em uma decisão teórico-metodológica crucial: focalizar os padrões coletivos de comportamento linguístico observáveis no uso concreto da língua.[2] Assim, o programa de pesquisa da Sociolinguística tem como objetivo central integrar os fatores sociais na análise de um estágio do desenvolvimento histórico de uma língua. Concentrando-se no fenômeno da mudança linguística, esse modelo desenvolveu toda uma teorização a partir da relação entre a variação linguística observada num determinado momento e a mudança em curso na estrutura da língua ao longo do tempo. O ponto de partida é a concepção de que a variação observada na fala de uma comunidade linguística não é aleatória mas determinada, tanto pela estrutura da língua, quanto pela estrutura social da comunidade. Na medida em que tal variação pode constituir um estágio de um processo de mudança em curso, a análise desse processo de mudança não pode se restringir aos fatores estruturais internos, devendo integrar também os chamados fatores externos, relativos à estrutura social.

---

[2] Cf. Labov (1982, p.18): "o objeto da descrição linguística é a gramática da comunidade de fala: o sistema de comunicação usado na interação social" (tradução do original).

E, com base nesse programa de pesquisa da mudança na estrutura da língua, a Sociolinguística propõe uma nova concepção para o objeto de estudo da linguística,[3] através de um **sistema heterogêneo** constituído por unidades e regras variáveis, formalização analítica para a gramática da comunidade de fala.[4] Assim sendo, fica claro o direcionamento da Sociolinguística para a dimensão sócio-histórica do fenômeno linguístico, ou, mais precisamente, para a constituição histórica da língua concebida como um fato social.[5] Nessa linha, o programa de pesquisa da Sociolinguística realizou profícuas investigações acerca da inserção da mudança linguística na estrutura social da comunidade de fala, revelando o(s) seu(s) ponto(s) de origem na estrutura social e suas vias de propagação através da identificação dos setores e situações que tendiam a impulsionar as mudanças e aqueles que tendiam a refreá-las. Abordou também, através da análise empírica dos juízos de avaliação subjetiva, o desenvolvimento da mudança no plano das relações ideológicas que se estabelecem no seio da comunidade de fala. Entretanto, o problema do encaixamento da mudança na estrutura linguística, malgrado alguns sucessos iniciais no nível fonológico, ainda permanece como um território em que a Sociolinguística por si mesma avançou muito pouco (LUCCHESI, 2004c, p.204-208). Isso se deve à ausência de uma teoria da estrutura produzida no âmbito da Sociolinguística, o que faz com que os seus pesquisadores recorram sobretudo aos esquemas funcionalistas, em suas análises do encaixamento dos processos de variação/mudança na estrutura da língua;[6] ou até, como fez o próprio Labov (1969), busquem no modelo da Gramática Gerativa os instrumentos para a sua análise linguística; uma conjugação que havia sido rejeitada pelo próprio Chomsky (1966).[7]

Essa incapacidade do programa de pesquisa da Sociolinguística em lidar com a estrutura da língua decorreria da sua incapacidade de integrar, na descrição do comportamento linguístico, o sistema da faculdade da linguagem, através do qual o ser humano dá uma forma verbal aos seus pensamentos. Como observou Romaine (1982, p.251), o problema está em estabelecer uma relação entre o

---

3 Diz ainda Labov (1982, p.21) que "a estratégia de WLH (1968) foi contribuir para uma teoria geral da linguagem através de uma teoria da mudança linguística" (tradução do original).

4 Cf. Labov (1982, p.17): "A condição normal da comunidade de fala é a da heterogeneidade: podemos esperar encontrar uma larga gama de variantes, estilos, dialetos e linguagens usados por seus membros. Mais ainda, esta heterogeneidade é parte integrante da economia linguística da comunidade, necessária para satisfazer às demandas linguísticas da vida quotidiana" (tradução do original).

5 Para uma definição das duas grandes dimensões do fenômeno linguístico, a sua dimensão sócio-histórica e a sua dimensão psíquico-biológica, veja-se Lucchesi, 2004c.

6 Cf. Poplack (1979, 1980a, 1981), Lefebvre (1981) e Scherre (1988), entre outros.

7 Embora Chomsky nunca tenha ignorado o papel do uso e da função no desenho da arquitetura da gramática: "Certamente é razoável supor que a necessidade de comunicação influencia a estrutura. Eu concordo" (CHOMSKY, 1975, p.56-58) (tradução do original).

comportamento linguístico dos indivíduos dentro de uma comunidade de fala e a sua competência linguística enquanto falantes nativos de uma dada língua natural:

> "Saber" inglês, por exemplo, não pode ser equiparado com saber as probabilidades ou hierarquias de efeito de diferentes contextos associados com regras na gramática. Nós podemos formular leis estatísticas sobre a probabilidade de ocorrências em uma língua e mesmo assim não sermos capazes de entender qualquer coisa que seja dita. Eu não me acho capaz de acreditar que o conhecimento de uma língua pode ser separado do papel que ele desempenha na compreensão. Eu não estou afirmando que Labov negue a existência desse tipo de (ou parte da) competência, mas simplesmente que não há qualquer coisa no conceito de regra variável que nos capacite a dizer coisas interessantes sobre esse aspecto da competência.[8]

Em outras palavras, a questão repousa na necessidade de integrar a dimensão psíquico-biológica do fenômeno linguístico em sua dimensão sócio-histórica, o que, em termos empíricos, significa definir em que medida os processos sociais de mudança linguística que ocorrem nas redes de relações sociais que definem a tessitura social da comunidade de fala são restringidos pelos mecanismos do sistema biológico da faculdade da linguagem, no sentido que tem sido definido por Chomsky (1986, 1995, entre outros). A percepção dessa lacuna no esquema de análise social da mudança linguística, em face de sua incapacidade de integrar o sistema mental de funcionamento da língua, está na base das significativas tentativas de conjugação dos modelos da **Teoria da Variação** e da **Teoria da Gramática** empreendidas por Tarallo (1987, 1991), Tarallo; Kato (1989), Kato (1999a), Ramos (1999), Duarte (1999), Kroch (1989, 2001, 2005), Kroch; Taylor (1997) e Roberts (2007).

Todas as tentativas de conjugar a abordagem gerativa e a sociolinguística têm de superar o desafio de reunir dois modelos que se fundamentam em concepções muito distintas do seu objeto de estudo. E, como as decisões teóricas e metodológicas fundamentais de um programa de pesquisa derivam axiomaticamente da concepção do seu **objeto de estudo**, tal obstáculo ainda se afigura como intransponível. Portanto, não se pretende aqui resolver uma questão de tal envergadura, mas esboçar uma visão que permita utilizar, de uma maneira teoricamente consistente, elementos da teoria da gramática na análise dos padrões coletivos de comportamento linguístico e vice-versa.

4.1.1 Estudo dos padrões coletivos de comportamento linguístico

A análise dos padrões coletivos de comportamento linguístico de uma dada comunidade de fala assenta no princípio teórico de que tais padrões são formalizados

---

[8] Traduzido do original.

analiticamente através de um sistema heterogêneo de regras e unidades variáveis, que são funções, tanto de fatores linguísticos, quanto de fatores sociais. Isso significa que os atos de fala que constituem os padrões de comportamento linguístico são regidos por fatores relativos:

(i) ao conhecimento intuitivo, que faz do indivíduo um membro da comunidade de fala (ou seja, um falante);

(ii) a certas características sociais desse indivíduo (tais como sexo, idade, nível de escolaridade, etc.);

(iii) aos juízos de valor desse indivíduo sobre as escolhas disponíveis no repertório linguístico da comunidade;

(iv) à dimensão ideológica do ato de interação verbal, que remete às relações de poder, estratégias de convencimento, nível de interesse na conversação, etc;

(v) às circunstâncias específicas nas quais se realiza o ato, relativamente ao conhecimento compartilhado entre os participantes, à intervenção eventual de participantes adventícios ao ato em questão, etc.

Buscando integrar a maior gama de fatores que possam ser apreendidos num esquema probabilístico através das variáveis arroladas na análise de um determinado fato da língua, tal análise poderá fazer predições em termos de agregados, e não de indivíduos.[9] Ou seja, não será capaz de afirmar categoricamente que o falante *X* de uma comunidade de fala *A* usará a forma linguística α no contexto linguístico /X__Y/, mas poderá afirmar que um falante qualquer com as características *x*, numa determinada situação *y*, etc, tenderá a usar a forma linguística α no contexto linguístico /X__Y/ na proporção *z*. Tal informação não é relevante para a compreensão do conhecimento internalizado que torna os indivíduos membros da comunidade de fala *A*,[10] mas constitui a base do conhecimento acerca do comportamento linguístico dos seus membros. Esse conhecimento dos padrões coletivos de comportamento linguístico tem uma importância intrínseca, enquanto conhecimento a que se deve chegar sobre uma determinada língua histórica, além da importância decorrente do fato de ser ele o terreno mais privilegiado para se tratar da questão da mudança linguística, na medida em que mudanças linguísticas e alterações nas frequências de uso estão intimamente relacionadas.[11]

---

[9] Romaine (1982, p. 244) observa que, no modelo da Sociolinguística Variacionista, "o local da língua é a comunidade ou o grupo, e a fala de qualquer grupo social será menos variável que a fala de qualquer indivíduo. Assim, as regras variáveis são escritas para grupos, não para indivíduos" (tradução do original).

[10] Cf. Romaine (1982, p. 251), citado acima.

[11] Cf. seção 4.1.2, sobre a relação entre as alterações de frequências e a mudança linguística, do ponto de vista das análises gerativistas.

A grande lacuna nesse modelo é a sua incapacidade em integrar o conhecimento linguístico internalizado na mente dos falantes que constituem a comunidade de fala. Nos termos da distinção chomskiana, ao formalizar analiticamente os padrões coletivos de comportamento linguístico, esse modelo pode dar conta da **língua-E**, mas não da **língua-I**. Adotando-se o pressuposto de que o comportamento linguístico dos indivíduos é, em grande medida, determinado pelo seu conhecimento linguístico internalizado, o poder explicativo do modelo variacionista fica bastante comprometido, se esse modelo não for capaz de integrar em sua explicação da mudança uma teoria forte sobre a estruturação do conhecimento linguístico na mente dos falantes.

Uma boa teoria acerca do conhecimento internalizado na mente dos falantes pode contribuir bastante para o enfrentamento de questões importantes para a compreensão da mudança linguística, e particularmente para uma questão crucial, não apenas para o entendimento da mudança mas para a compreensão do fenômeno linguístico como um todo: *quais estruturas da gramática de uma comunidade de fala são passíveis de variação e quais não o são?* Em outras palavras, coloca-se aqui a questão de estabelecer **os limites da variação na estrutura da língua** — uma questão que ainda ocupa uma posição marginal entre os paradigmas vigentes na linguística contemporânea, mas cujo enfrentamento pode determinar um salto de qualidade na compreensão do funcionamento da linguagem humana. A resposta a esse tipo de questão é crucial para se definir quais tipos de escolhas estão potencialmente disponíveis no uso da língua. Desse modo, a teoria da gramática é de fundamental importância, não apenas para explicar a implementação de um processo de mudança, mas sobretudo o seu ponto de partida (TARALLO, 1991), definindo os processos potenciais de mudança em uma determinada língua particular, ou nas línguas humanas em geral, no que se designou no âmbito da Teoria Sociolinguística como **problema das restrições** (ing. *constraints problem*).

Por outro lado, no processo de difusão da mudança, a estrutura da gramática vai determinar também quais os contextos que mais favorecem uma determinada variante e quais os que favorecem outra. Há de ter-se em mente que, no jogo cambiante de processos de variação na estrutura da língua, a alteração das frequências de uso de uma variável pode ter implicações diretas e decisivas no desenvolvimento diacrônico de uma outra variável linguística, o que se situa no plano do **problema do encaixamento** (ing. *embedding problem*). Portanto, estabelecer o que leva à fixação de uma determinada variante em detrimento das demais, no âmbito de um processo histórico de mudança linguística (o **problema da implementação** — ing. *actuation problem*), depende tanto de uma acurada observação dos fatores sociais, culturais e ideológicos envolvidos no processo, quanto de uma boa compreensão do funcionamento da gramática.

Assim sendo, a possibilidade de integração de um modelo poderoso de análise da gramática, como o modelo gerativista, no estudo dos padrões coletivos de comportamento linguístico e, consequentemente, no estudo das mudanças linguísticas e de outros processos históricos, tais como os processos de **transmissão linguística irregular**, torna-se uma opção deveras aliciante. O problema teórico fundamental com que se depara tal intento está no fato de a Gramática Gerativa formalizar o conhecimento linguístico internalizado na mente do falante, a língua-I, em termos de unidades invariáveis e regras categóricas, enquanto os padrões coletivos de comportamento linguístico são formalizados analiticamente, no âmbito da Sociolinguística Variacionista, por meio de unidades e regras variáveis.

A gramática de uma língua (isto é, a formalização analítica da língua-I) constitui a resposta que o Programa Gerativista apresenta para uma questão central da investigação linguística: *como a língua funciona?* Situando o problema no conhecimento internalizado na mente do falante, e concebendo esse conhecimento como resultante do desenvolvimento da faculdade inata da linguagem, o Programa Gerativista situa claramente a sua abordagem na dimensão psíquico-biológica do fenômeno linguístico. Ao se concentrar sobre a relação entre variação e mudança, a Sociolinguística Variacionista situa-se na dimensão sócio-histórica do fenômeno linguístico, devendo, portanto, operar com esquemas probabilísticos de análise, mais adequados aos processos históricos, que são, em grande medida, particulares e contingenciais, em função da combinação específica e fortuita dos fatores que os determinam. Superar esse imenso fosso teórico e metodológico é o grande desafio para aqueles que buscam conjugar a teoria da gramática e a análise sócio-histórica da mudança linguística.

4.1.2 A mudança linguística e a teoria da gramática gerativa

A abordagem da mudança linguística assumida pelos gerativistas se baseia nas seguintes assunções fundamentais: a criança constrói sua gramática a partir do que ouve (os dados linguísticos primários — PLDs) e do que possui de inato (os princípios da gramática universal — UG). Como a criança não tem acesso direto à gramática das pessoas à sua volta, as conclusões a que ela chega sobre sua gramática são baseadas inteiramente nessas duas fontes do conhecimento linguístico: os dados primários e os princípios universais. Portanto, a aquisição é vista como um processo de seleção de uma gramática a partir dos PLDs e dos princípios da UG.[12]

[12] As questões aqui discutidas sobre mudança linguística na perspectiva da Gramática Gerativa se apoiam fundamentalmente nos estudos de Lightfoot (1979, 1991 1999b); Roberts (1993, 2007); Kroch (1989, 2001, 2005). Ressalta-se que a Gramática Gerativa não desenvolveu uma teoria da mudança linguística independente da teoria da gramática, como o fez a Sociolinguística.

A Gramática Universal, a teoria dos universais linguísticos biologicamente determinados, é uma caracterização abstrata da noção de língua humana possível, e tenta determinar o que pode variar entre as línguas (os parâmetros) e o que permanece constante (os princípios). Os princípios são representações abstratas dos mecanismos mentais usados pelos falantes quando processam uma sentença, ou seja, são hipóteses empíricas dos mecanismos inatos predeterminados, que podem ser falseadas ou analisadas como inadequadas a partir da descoberta de estruturas linguísticas que as violem sistematicamente. Os parâmetros definem o espaço das variações entre as línguas, variações tais determinadas por um conjunto de opções definido pela UG. Os valores designados a um dado parâmetro constituem, assim, as diferenças primitivas entre dois sistemas gramaticais; tais diferenças primitivas interagem com a estrutura dedutiva da UG, que permanece constante, e com outros parâmetros, uma interação que determina muitas diferenças derivadas.

Embora as propriedades da UG permaneçam constantes de geração em geração, os PLDs mudam, e, assim, o *input* para uma geração pode não ser o mesmo para a próxima. Como a gramática a ser adquirida é uma consequência da interação entre propriedades da UG biologicamente determinadas e o ambiente — e o ambiente é constantemente afetado, de muitas maneiras, por mudanças culturais, sociais e pragmáticas, que se tornarão parte do *input* linguístico para uma próxima geração —, a mudança é inevitável. Portanto, é no processo de aquisição da linguagem, visto como resultado de um processo abdutivo (lei + resultado => caso)[13] que a mudança pode ocorrer (LIGHTFOOT, 1979).

A abordagem paramétrica se propôs, inicialmente, a explicar as diferenças sintáticas entre as línguas numa dimensão sincrônica. O desenvolvimento para a dimensão diacrônica é bastante natural, uma vez que as diferenças sintáticas entre dois ou mais estágios de uma língua são analisadas em termos de diferentes valores paramétricos, do mesmo modo que se analisam as diferenças sintáticas entre as línguas contemporâneas. Assim, na perspectiva da **Teoria de Princípios e Parâmetros**, a maneira natural de olhar uma mudança sintática é como uma **mudança no valor de parâmetro**.

Os estudos sobre diferenças paramétricas entre duas ou mais línguas diferentes procuram descrever o parâmetro e suas consequências, as suas implicações tipológicas e as experiências desencadeadoras subjacentes à seleção desses valores paramétricos; os estudos sobre diferenças paramétricas entre dois estágios de uma mesma língua histórica precisam explicar não apenas qual é o parâmetro em questão, quais são os seus efeitos, mas também *como, em algum ponto da transmissão da*

[13] Cf. seção 4.2, para algumas questões pontuadas sobre o processo de aquisição de L1, de L2 e em situações de contato.

língua de uma geração para outra, o novo valor foi favorecido em detrimento do outro (ADAMS, 1987; VANCE, 1989; ROBERTS, 2007).

Três aspectos fundamentais da relação entre a visão geral de parâmetro, de aquisição e de mudança têm sido destacados nas discussões teóricas: (i) o da natureza discreta dos valores dos parâmetros; (ii) o de que a mudança é catastrófica, no sentido de que a mudança do valor de um parâmetro é abrupta e irreversível num dado momento histórico; (iii) o de que a mudança é interna, ou seja, é um fenômeno de língua-I, em princípio independente do ambiente social, cultural e histórico do aprendiz. Esses três aspectos parecem entrar em contradição com o que estudos de variação e mudança linguísticas têm assumido, implícita ou explicitamente: a gradualidade da mudança linguística e a centralidade de fatores externos de vários tipos para entender a natureza da mudança. Contudo, essa contradição é aparente, desde que se consiga seguir a linha de raciocínio de cada uma das duas correntes teóricas aqui em discussão, ou seja, desde que se procure ver como reconciliar as ideias de sistemas algorítmicos homogêneos, discretos, que são propriedades dos indivíduos, com a mudança gradual da linguagem num dado momento histórico em uma dada comunidade de fala, formalizada analiticamente, pelos sociolinguistas, por algorítmicos heterogêneos, graduais e probabilísticos.[14]

A razão principal para olhar a mudança como gradual centra-se nos próprios estudos de mudança, que traçam uma evolução temporal de competição entre formas conservadoras e inovadoras, como observadas nos registros históricos (além das ideias de evolução biológica que influenciaram a linguística histórica desde o século XIX). Outra razão fundamental é que a mudança linguística não é um evento histórico com data marcada específica; ao contrário, parece ser temporalmente difusa. Há também razões conceituais para ver a mudança como catastrófica, abrupta, mesmo quando se consideram não só fatores gramaticais como também os sociolinguísticos.

A primeira das razões conceituais está explícita na seguinte citação de Lightfoot (1979, p.391), que focaliza o problema sem solução da noção de deriva tipológica, pois as crianças não podem abstrair informações de mudanças em curso nos PLDs a que estão expostas:

> Línguas são adquiridas e gramáticas são construídas por indivíduos de cada geração. Eles não têm memórias raciais de forma a saber, em qualquer sentido, que suas línguas estão gradualmente se desenvolvendo de um tipo SOV para um tipo SVO, por exemplo, e que elas devem

[14] Seguindo-se para isso as proposições delineadas por Kroch (2001) e Roberts (2007).

> continuar nesta direção. Além disso, se houvesse uma hierarquia determinada de mudanças a acontecer, como poderia uma criança, confrontada com uma língua exatamente a meio caminho da hierarquia, saber se a língua estava mudando do tipo X para o tipo Y ou vice-versa?[15]

A segunda das razões é uma questão lógica, a *Lei do Meio Excluído* (ing. *Law of the Excluded Middle*: p v p) (ROBERTS, 2007, p.296). Portanto, tratar valores paramétricos como entidades formais discretas (entidades como, por exemplo, verbo, nome, tempo, fonema, etc.) implica que determinados traços formais têm um valor ou outro, mas não podem estar no meio (X não pode ser +/- um fonema de uma língua), ou seja, nenhum sistema pode estar em um estágio intermediário entre dois valores paramétricos. Por isso Kroch (1989) ressalta que a mudança de uma geração para outra é necessariamente instantânea.

Roberts (2007) discute como se pode conciliar o curso temporal da mudança linguística, como observada nos dados históricos, e o fato de que a frequência de ocorrências entre formas inovadoras e conservadoras segue uma curva-S, analisando o efeito da gradualidade na perspectiva de fatores sociais (tipos de registro, variação individual, idade, etc.) e na da natureza do sistema gramatical (difusão lexical, microparâmetros, verdadeira opcionalidade na gramática, macroparâmetros). A combinação desses fatores gramaticais e sociais, independentes um do outro, é responsável pela *miragem* da mudança gradual, escondendo, nos dados históricos, os efeitos da mudança estrutural discreta e instantânea.

A difusão lexical, ou seja, a ideia de que mudanças que afetam o léxico podem se difundir gradualmente, tem sido bastante atestada em mudanças sonoras (LABOV, 1994) e em mudanças de traços formais do léxico substantivo. Um bom exemplo do último tipo é a que aconteceu com os verbos psicológicos entre o inglês antigo e o inglês médio. Lightfoot (1991) observa que, no inglês antigo, a forma correspondente do verbo *like* (*lician*) se realizava usualmente em construções com um NP experienciador no caso dativo inerente e um NP tema no nominativo; a perda do caso oblíquo morfológico no inglês antigo vai se refletir em que um NP experienciador dativo não poderia mais ser produzido, tendo sido reanalisado como um sujeito nominativo, o NP tema passando a receber caso acusativo. Assim, uma construção do tipo *Ao rei agrada a vitória* passa a ser realizada como *O rei gosta/ama a vitória*.

Contudo, esse tipo de mudança afeta somente entradas lexicais de verbos individuais, difundindo-se sobre a classe de verbos psicológicos durante um período de tempo, num processo de reanálise localizada em uma parte das estruturas apropriadas a esses verbos, mas isso não é uma mudança paramétrica. Desde a proposta em Borer (1984), a mudança paramétrica tem sido vista como alteração

[15] Tradução do original.

no conjunto de traços do léxico gramatical, funcional, e não no do léxico substantivo. Mudanças deste último tipo afetam estruturas específicas, mas não a gramática como um todo (ROBERTS, 1993, 2007); mudanças do primeiro tipo, alterando o valor de um dos traços do sistema de tempo de uma língua, por exemplo, afetará todos os tipos de sentenças finitas, e essa mudança tem de ser necessariamente discreta, abrupta, instantânea no indivíduo. Essa distinção entre **mudanças paramétricas** e **mudanças no léxico substantivo** é, portanto, crucial para a formulação de uma taxonomia mais abrangente da mudança linguística que permita deslindar processos de mudança graduais de processos de mudança abruptos na estrutura da língua e capturar, dessa forma, especificidades de cada processo sócio-histórico, com um valor explicativo mais definido e, portanto, mais profícuo.

As categorias funcionais são compostas de um conjunto de traços formais. Segundo Roberts (2007), isso abre a possibilidade de difusão lexical através do sistema funcional, quando uma série de mudanças discretas atinge gradualmente alguns traços do conjunto de traços de categorias funcionais.[16] Essa ideia tem sido desenvolvida a partir da noção de **microparâmetros**, que procura identificar as unidades mínimas de variação paramétrica em dialetos contemporâneos de uma mesma língua histórica (KAYNE, 2000). Um bom exemplo de variação microparamétrica é observado na seleção de verbo auxiliar na voz ativa, em dialetos do italiano, entre os correspondentes a *haver* e *ser*: no italiano *standard*, a seleção observa a estrutura argumental do verbo lexical, e *ser* só é selecionado nas construções inacusativas; em dialetos das áreas central e sul da Itália, a seleção leva em conta a especificação de pessoa do sujeito, sendo *ser* a forma selecionada para 1ª e 2ª pessoas, independentemente da estrutura argumental do verbo; na área urbana napolitana, só a forma correspondente a *haver* é usada.[17] A variação observada no conjunto pode apontar para uma mudança gradual nas comunidades, mudança que pode ser reduzida à escolha de valores discretos das categorias funcionais em que os auxiliares são inseridos. Analisando esses fatos, Roberts (2007, p.300) conclui que:

> A possibilidade de difusão lexical de tais tipos de traços cria a expectativa de podermos observar uma variação sincrônica muito sutil, devida à variação em diferenças mínimas dos traços de núcleos funcionais, e uma mudança diacrônica aparentemente gradual, à medida que esses traços alteram seus valores.[18]

---

[16] Esta questão será retomada na seção 4.2.

[17] Ladgeway (2000) apresenta outras possibilidades em outros dialetos do italiano. Roberts (2007, p.300-305) apresenta uma explicação formal para esta microvariação.

[18] Tradução do original.

Biberauer e Roberts (2005) e Roberts (2007) mostram que, mesmo no modelo minimalista da gramática gerativa, opcionalidade formal é uma possibilidade técnica, como Biberauer e Richards (2006) demonstraram com relação ao movimento-WH em russo, em que o movimento *pied piping* é opcional.[19] No caso de opcionalidade formal, as duas formas podem gradualmente variar em frequência no curso do tempo, revelando uma mudança gradual nas escolhas de natureza sociolinguística, sobretudo no valor social que uma das formas pode adquirir, levando a seu uso predominante, mas isso não constitui uma mudança paramétrica. Ao assumir essa opcionalidade formal em uma mesma gramática, ou seja, uma variação estrutural sem consequências semântico-pragmáticas, alcança-se uma base teórica formal para se enfrentar a complexidade da variação nos usos dentro de uma mesma comunidade de fala.

Kroch (1989), apoiado em um conjunto de fenômenos de mudança, analisadas com recurso a uma ferramenta estatística sofisticada, apontou a possibilidade de olhar a gradualidade da mudança em termos de gramáticas em competição. Segundo ele, a mudança paramétrica tende a apresentar uma curva-S, quando as frequências de formas antigas e novas são comparadas temporalmente, e esta curva-S pode ser matematicamente modelada. Kroch (1989, p.200) diz que "quando uma opção gramatical substitui outra com a qual está em competição em um conjunto de contextos linguísticos, a frequência dessa substituição, medida adequadamente, será a mesma em todos esses contextos",[20] configurando o que ele denominou de **Efeito de Taxa Constante** (ing. *Constant Rate Effect*).[21] A proposta de Kroch levanta duas questões interessantes para o estudo da mudança. Primeiro, permite que a mudança gradual possa ser vista como competição entre entidades distintas, ou seja, como uma mudança paramétrica vai se implementando na comunidade de fala; segundo, a curva-S não deve ser considerada um fato da gramática em si mesma, e sim um fato sobre a forma como os indivíduos com competência em mais de uma gramática acessam as opções disponíveis e sobre como os fatores sociais atuam nessas escolhas.

Embora as abordagens imanentistas expliquem a mudança em nível individual, nada dizem sobre como a mudança afeta comunidades de fala, questão central na abordagem sociolinguística. Nos termos dessa abordagem, estas duas questões podem ser integradas no **problema da transição** da mudança: como a nova forma se difunde na comunidade de fala? Já a proposta gerativista de que a

---

[19] De uma forma bem simplificada, o fenômeno pode ser esquematizado da seguinte maneira: [*o livro de quem*]$_i$ *você leu t*$_i$*?* / [*de quem*]$_i$ *você leu* [*o livro t*$_i$]?)

[20] Tradução do original.

[21] Kroch (1989) ilustra o *Efeito de Taxa Constante* com um número de fatos da sintaxe diacrônica do inglês, relacionados com o valor negativo do movimento do verbo para Tempo. Cf. também nota 28.

mudança linguística é direcionada pela aquisição da linguagem fornece resposta para o **problema da implementação**.[22] A noção de gramática em competição, de Kroch, e a hipótese do *Efeito de Taxa Constante* podem explicar o problema da transição,[23] a difusão da nova gramática nos indivíduos e na comunidade de fala, considerando a existência de indivíduos e comunidades de fala bilíngues, quando mais de uma gramática (no sentido da *competência linguística*) está presente numa mesma comunidade de fala em um determinado momento histórico.

A noção de competição de gramática de Kroch (2001, p.723) pode ser resumida na seguinte citação:

> Formas em competição em uma língua, em situação de diglossia sintática, podem facilmente representar uma opção entre uma inovação vernacular e um uso literário conservador. Uma vez que a forma inovadora tenha vantagens em termos psicológicos e de número de falantes, poderá ao longo do tempo vencer a competição, mesmo em textos escritos. Com base nesse modelo, a gradualidade encontrada em textos não deve refletir qualquer mecanismo básico de mudança linguística, e sim mecanismos psico- e sociolinguísticos de bilinguismo.[24]

Assim, as forças externas que atuam na competição entre gramáticas refletem-se nas variações associadas a valores sociais relacionados com a diglossia sintática (nível formal, coloquial, etc.) e com o bilinguismo, que permitem fenômenos de *code-switching*, *code-mixing*, amplamente atestados em comunidades bilíngues.[25]

Roberts (2007) diz que o conceito de **diferenciação ordenada** (ing. *orderly differentiation*) na estrutura social, noção chave para entender a mudança linguística dentro do modelo da Teoria da Variação (WLH, 1968 [1989, p.151]), não pode ser explicado só pela competição entre gramáticas em si mesmas, porque, se assim fosse, se estaria negando a possibilidade real de falantes bilíngues em línguas nativas. A competição de gramática não explica, nem o início da mudança (*constraints problem*), nem sua conclusão (*actuation problem*); o primeiro relativo às possibilidades estruturais potencialmente disponíveis para o uso linguístico, e o segundo identificado com o sucesso da gramática inovadora sobre a conservadora. A questão

---

[22] Embora WLH (1968 [1989, p.145-146]) critiquem uma proposta similar apresentada por Halle (1962). Veja-se ainda a seção 4.1.3 para uma discussão mais detalhada dos problemas da mudança, nas duas abordagens teóricas em destaque.

[23] Ou seja, "os estágios intervenientes que podem ser observados, ou que devem ser postulados, entre quaisquer duas formas de uma língua definida para uma comunidade linguística em diferentes períodos" (WLH, 1968 [1989, p.101]). Observa-se que o *Constant Rate Effect* é um aspecto da transição da mudança de um ponto zero para valores positivos temporalmente crescentes (KROCH, 1989, p.205).

[24] Tradução livre do original.

[25] Cf. seção 4.2.

colocada por Roberts (2007, p.320) é: por que as duas gramáticas não convivem em paz? por que uma suplanta a outra? Isso também está relacionado à síntese apresentada por WLH (1968 [1989, p.184-185]), sobre o *transition problem*:

> Esta transição ou transferência de traços de um falante para outro parece acontecer por meio de falantes bidialetais ou mais geralmente, falantes com sistemas heterogêneos caracterizados pela diferenciação ordenada. Mudanças acontecem (1) quando um falante aprende uma forma alternativa, (2) durante o tempo em que as duas formas existem em contato na sua competência, e (3) quando uma das formas se torna obsoleta.[26]

As questões de mudança sintática pontuadas acima recebem uma explicação no estudo de Roberts (2007, p.334-335), com base no modelo de mudança paramétrica, nos seguintes termos. As **formas alternativas** podem indicar ou opções geradas por um único sistema ou competição entre gramáticas na competência linguística de falantes de uma comunidade. Uma nova opção paramétrica, ou seja, uma nova gramática ou a remarcação do valor de um parâmetro de modo que uma nova opção é gerada, acontece no estágio (1), alterando a proporção de indivíduos bilíngues (com duas gramáticas) ou do valor de um parâmetro permitindo opcionalidade formal; a opção paramétrica antiga ou a perda de uma opção formal desaparece no estágio (3). O estágio (2) reflete o modo como o novo parâmetro substitui o antigo, ou seja, como o novo valor paramétrico vai se difundir na comunidade, mudando gradualmente a proporção de indivíduos em que as duas formas coexistem, através da competição de gramáticas ou da opcionalidade formal gerada por um único sistema.

Os gerativistas procuram desenvolver um modelo para a difusão gradual da mudança através de uma comunidade de fala centrado na visão de que a implementação da mudança está relacionada com questões de aquisição da linguagem e de que a noção de parâmetros binários e discretos define o local de variação nos sistemas linguísticos. As relações entre aquisição, mudança e alterações nas frequências de uso podem ser sintetizadas como segue. A mudança linguística acontece em consequência de "erros" no processo de aquisição, de L1 ou de L2; muitos desses "erros" são compatíveis com a GU e podem se propagar para todos os membros da comunidade (ou regredir), a depender das evidências que a língua-alvo oferece, do valor estrutural marcado ou não marcado da forma inovadora, bem como dos valores sociais que adquirem na comunidade. Essa é a posição assumida por Kroch (2005), para quem a mudança no valor de um parâmetro precede as alterações nas frequências de uso; tais alterações refletem uma situação de bidialectalismo. Outros linguistas consideram que alterações nas frequências de

[26] Tradução do original.

uso, por questões de estilo ou estratégias discursivas, quando um tipo particular de construção pode se tornar mais frequente como resultado de ter adquirido alguma função expressiva, podem tornar opaco o valor de um parâmetro, levando a criança a optar por um valor não marcado do parâmetro. Nesse caso, as alterações nas frequências de uso precedem a mudança paramétrica. Essa é a explicação proposta para a perda da restrição V2 do francês medieval, em que se tem explicitamente assumido como evidência crucial da perda dessa propriedade o número crescente de construções V>2 (de deslocadas à esquerda) atestadas, sobretudo, nos documentos franceses do século XV (ADAMS, 1987; VANCE, 1989).

Desse modo, a quantificação tem sido usada como um guia útil na confirmação das tendências gerais do fenômeno sintático em estudo e da sintaxe do texto. Tem-se considerado que a frequência com que diferentes padrões de realização estrutural de constituintes ocorrem permite fazer certas generalizações sobre a sintaxe dos documentos em questão. Considera-se, por exemplo, que, se uma determinada ordenação de palavras é rara e, sobretudo se ela não se ajusta ao amplo padrão sintático do texto, é muito provavelmente um arcaísmo ou artifício literário (VANCE, 1989). Assim, a frequência dos dados pode certamente esclarecer sobre a validade das hipóteses. Contudo, deve-se ter cuidado com generalizações desse tipo e com as inferências derivadas de frequências, desde que a não ocorrência de uma determinada ordem pode ser em razão de fatores ocasionais de limites do *corpus*; assim, é preciso muito mais do que a frequência de uma construção para motivar a escolha de uma estrutura de base (DUPUIS, 1989, p.5).

Os gerativistas procuram resolver as questões acima através da sintaxe comparada. O primeiro passo da abordagem gerativista nos estudos diacrônicos é o de depreender uma gramática coerente do fenômeno sintático em estudo nos textos selecionados. Escrever uma gramática de uma língua não mais falada envolve uma tarefa que não difere de qualquer outro estudo sincrônico, exceto pelo fato de que os dados são finitos e limitados a fontes escritas. A partir do momento em que se tem consciência da limitação dos dados, os métodos da teoria sintática gerativa servem bem a esse tipo de investigação.[27] O segundo passo (a comparação com outros sistemas linguísticos) beneficia-se particularmente do sistema de princípios e parâmetros da teoria gerativa. Assumindo que as línguas variam parametricamente, as diferenças observadas entre, por exemplo, o português arcaico, o português clássico e o português brasileiro podem ser vistas como o resultado da remarcação de um ou mais parâmetros durante o seu desenvolvimento histórico.

Em síntese, conjugar análises quantitativas e análises teóricas é uma metodologia profícua ao entendimento do fenômeno da mudança linguística, cada uma

[27] Desconsiderando aqui os métodos usados pelos gerativistas para obter dados negativos.

delas fornecendo ferramentas importantes que permitem evitar conclusões inadequadas sobre as características estruturais dos fatos em estudo e seus desenvolvimentos diacrônicos.

4.1.3 Síntese dos desenvolvimentos recentes e as alternativas que se colocam para a conjunção dos dois modelos

Os problemas da mudança linguística do ponto de vista da Sociolinguística, como discutidos em WLH (1968 [1989]), também são problemas fundamentais na pesquisa da Gramática Gerativa, embora sejam formulados diferentemente, em termos do modo como se entende o processo de aquisição da linguagem e as restrições impostas pela GU. O **problema da implementação**, o porquê do começo de uma mudança linguística, encontra uma resposta no programa gerativista em termos da cadeia de aquisição:

(i) a aquisição resulta de um processo seletivo e abdutivo, de interação entre estruturas cognitivas inatas e tipos de experiência selecionados;

(ii) os dados da experiência relevantes para a aquisição são constituídos de elementos robustos e estruturalmente simples;

(iii) a criança é um aprendiz grau-0 — sensível somente a dados de sentenças raiz e do complementador da sentença encaixada;

(iv) a morfologia desempenha papel importante na definição do valor de um parâmetro;

(v) parâmetros são marcados nos estágios iniciais da aquisição.

Assim, a mudança sintática surge quando acontece reanálise abdutiva de partes dos PLDs, reanálise que tem como causas processos morfofonológicos, ambiguidade estrutural de dados dos PLDs na expressão do valor paramétrico, inovações nos PLDs, situações de contato linguístico, fatores extralinguísticos, dentre outras (LIGHTFOOT, 1991; ROBERTS, 2007; KROCH, 2005).

O **problema do encaixamento** da mudança é analisado pelos sociolinguistas em duas direções: como a mudança está encaixada na estrutura linguística e na estrutura social. No âmbito da Gramática Gerativa, o encaixamento na estrutura linguística pode ser visto de três formas. Primeiro, a noção de parâmetro relaciona valores positivos ou negativos dos parâmetros a certas possibilidades estruturais, mas não à gramática como um todo, posição também assumida por WLH (1968 [1989, p.185]), quando afirmam que "a mudança linguística em si mesma raramente resulta no movimento de um sistema inteiro em outro sistema". Assim, as propriedades que podem variar entre as línguas não são autônomas; ao contrário, elas se organizam em grupos razoavelmente estáveis que permitem a identificação

de tipos de língua. Isso significa que, se uma língua tem uma propriedade A, ou seja, o valor positivo para um dado parâmetro X, ela também terá as propriedades B e C, mas não as D e E. Desse modo, a abordagem paramétrica pode explicar, em alguns casos, por que certas propriedades coocorrem durante a mudança.

A correlação entre parâmetros e propriedades estruturais pode ser observada em diferentes estudos de sintaxe diacrônica. Por exemplo, o Parâmetro do Sujeito Nulo[28] procura definir um número de propriedades que diferencia sistematicamente línguas como o português europeu, o italiano e o espanhol de línguas como o inglês e o francês (RIZZI, 1982): (a) possibilidade de sujeito pronominal referencial, definido, sem realização fonológica, em sentenças finitas discursivamente neutras; (b) inversão entre sujeito e verbo (ordem VS) com diferentes tipos de verbo; (c) ausência do efeito do filtro *that-trace*; (d) possibilidade de *clitic climbing* (KAYNE, 1989); (e) ausência de sujeito expletivo com verbos que exprimem fenômenos da natureza. Considera-se, assim, que, se uma língua perde a propriedade básica correspondente ao valor do parâmetro, também perderá as outras propriedades relacionadas com o valor do parâmetro.[29]

Os estudos de Lightfoot (1979, 1991) sobre a reanálise dos verbos plenos *can, could, do, did, may, might, must, shall, should, will, would* como verbos auxiliares, do inglês antigo ao inglês médio, apresentam um outro exemplo de encaixamento estrutural: (a) esses verbos perderam a habilidade de selecionar um objeto direto; (b) tornaram-se flexionalmente distintos dos outros verbos na 3ª pessoa do singular; (c) a relação entre tempo presente e passado tornou-se de alguma forma não temporal (indicam modo subjuntivo ou condicional); (d) não aceitam serem seguidos por infinitivas introduzidas por *to*. Assim, a reanálise desses verbos como modais foi desencadeada por um conjunto de mudanças gramaticais distintas.

Roberts (2007, p.355) chama atenção para o fato de que uma mudança no sistema funcional de uma língua pode desencadear mudanças paramétricas em cascata (efeito dominó), fazendo com que uma língua mude de um tipo para outro. Isso parece ser o que aconteceu com o inglês, entre 1100 e 1700, passando a divergir radicalmente de outras línguas germânicas, como também das línguas românicas: (i) perda de V2 (1450) > (ii) lexicalização de Tempo com auxiliares modais e a forma *do* (1525) > (iii) perda de movimento do verbo para Tempo (1575)

---

[28] Vale a pena ressaltar que estão sendo apresentadas propriedades amplas do parâmetro. Uma análise mais adequada da questão tem de considerar, sobretudo, qual é a propriedade específica da língua que licencia qual tipo de sujeito nulo, pois os estudos têm mostrado que há diferenças relevantes entre as línguas de sujeito nulo consistente (italiano, espanhol), línguas de sujeito nulo parcial (finlandês, hebraico) e línguas de sujeito nulo radical (chinês, japonês) (HOLMBERG, 2004, 2005).

[29] O *Efeito de Taxa Constante* proposto por Kroch prediz que a difusão da mudança deve ser uniforme nos diferentes contextos relacionados com o valor paramétrico.

> (iv) contração da negação (1600) > (v) desenvolvimento de auxiliares negativos (1630) > (vi) desenvolvimento de *do* suporte (finais do século XVII). Essas mudanças em série, que afetam os traços de um único núcleo funcional, levaram a uma reorganização ampla no sistema de auxiliar e posicionamento de verbo e criaram "um sistema que é bastante diferente de qualquer um encontrado em línguas germânicas ou românicas" (p.355).

A questão do **encaixamento social**, bastante escrutinada no campo de pesquisa da Sociolinguística, também desempenha um papel crucial no algoritmo gerativo da mudança linguística, desde que mudanças no comportamento linguístico dos falantes adultos alteram os PLDs disponíveis para a aquisição da língua materna de uma geração a outra. Nesse sentido, a implementação de uma mudança por reanálise abdutiva pode ser desencadeada pela alteração nas frequências de uso de formas específicas, o que vai alterar os dados robustos disponíveis para aquisição da língua materna de uma geração em relação à precedente. Assim, como observa Lightfoot (1993, p.288), "a questão, então, passa a ser: por que a experiência desencadeadora na infância mudou para as duas gerações?".

Lightfoot (1993, p.290), assim como os gerativistas em geral, assume uma atitude um tanto quanto cética diante da possibilidade de resposta a tal questão:

O que em geral não podemos explicar é por que o ambiente linguístico deveria ter mudado inicialmente. As mudanças ambientais são frequentemente motivadas pelo que tenho chamado de fatores "do acaso", efeitos de empréstimos, mudanças na frequência de formas, inovações estilísticas, que se espalham por uma comunidade e, quando temos sorte, são documentadas por estudos de variação.

Revela-se, neste ponto, a complementaridade entre os dois modelos. Os fatores "do acaso" de Lightfoot nada mais são do que os fatores sociais que condicionam as escolhas que os falantes fazem no plano concreto e específico do uso linguístico. Ao descrever as mudanças no plano social e cultural, que determinaram as mudanças nos padrões de comportamento linguístico no seio de uma determinada comunidade de fala, a análise sociolinguística poderá definir os cenários subjacentes a uma determinada mudança na gramática. Dessa forma, fica clara também a distinção proposta por Lightfoot (1993, p.291) entre as mudanças "que se devem ao acaso" e aquelas que se devem "à necessidade, ou seja, novas fixações de parâmetro". Como mudanças do primeiro caso, tem-se, na história do português, a substituição dos pronomes pessoais canônicos *tu/vós* e *nós* em função da gramaticalização das formas nominais *você(s)* (< *vossa mercê*) e *a gente*, respectivamente. Não se trata de uma mudança determinada pela força da gramática, mas de uma mudança de hábitos linguísticos correlacionada a mudanças comportamentais no plano da cultura, da ideologia e das relações sociais.

Contudo, tais mudanças têm implicações gramaticais, na medida em que reduzem o espectro flexional do verbo. A assunção de *você(s)* e *a gente* reduz o antigo sistema de seis formas distintas do verbo a apenas três, ou mesmo a duas, como se pode ver no Quadro 1.

PRESENTE DO INDICATIVO		PRETÉRITO IMPERFEITO DO INDICATIVO	
Eu trabalh**o**	Eu trabalh**o**	Eu trabalhava	Eu trabalhava
Tu trabalha**s**	Você trabalha	Tu trabalhava**s**	Você trabalhava
Ele trabalha	Ele trabalha	Ele trabalhava	Ele trabalhava
Nós trabalha**mos**	A gente trabalha	Nós trabalháva**mos**	A gente trabalhava
Vós trabalha**is**	Vocês trabalha**m**	Vós trabalháve**is**	Vocês trabalhava**m**
Eles trabalha**m**	Eles trabalha**m**	Eles trabalhava**m**	Eles trabalhava**m**

Quadro 1 - Flexão verbal de número e pessoa com a introdução de *você(s)* e *a gente* na pauta dos pronomes pessoais

Essa mudança no ambiente linguístico, devida ao "acaso" das alterações sócio-históricas no plano do comportamento linguístico dos falantes adultos, pode criar as condições propícias para uma mudança necessária, ou seja, uma mudança na gramática; mais precisamente, uma alteração no parâmetro do sujeito nulo no português brasileiro. Uma hipótese aliciante é a de que um avanço maciço de *a gente* sobre *nós*, tornando residuais as formas verbais flexionadas em *-mos*, poderá criar as condições necessárias para uma mudança paramétrica, de modo que o sujeito nulo deixe de ser licenciado no português brasileiro. Desse modo, tem-se, por um lado, uma mudança comportamental (a substituição de *nós* por *a gente*), que está correlacionada a alterações no plano da cultura, da ideologia e das relações sociais. Por outro lado, tem-se uma mudança gramatical: uma alteração no valor do parâmetro do sujeito nulo. Esta é determinada pela mudança no ambiente linguístico que fornece os PLDs para aquisição da língua materna; como hipótese, a obsolescência das formas em *-mos* eliminará a possibilidade de dados robustos que licenciem o sujeito nulo. Enquanto a primeira é gradual — a variação estruturada entre *nós* e *a gente* já tem mais de duzentos anos (LOPES, 1999) — e está à mercê das contingências do devir sócio-histórico, a última é necessária, desde que ocorram as devidas alterações no ambiente linguístico, e abrupta (devendo-se completar no período da emergência de uma geração de falantes cuja gramática não licenciará mais o sujeito nulo). Tal taxonomia da mudança linguística abre caminhos para superar o impasse entre a visão da mudança linguística socialmente gradual e a concepção da mudança linguística abrupta e catastrófica no plano da gramática.

Por outro lado, a concepção da concorrência entre gramáticas distintas em uma mesma comunidade de fala, como proposta por Kroch (2001, 2005), permite conciliar a concepção de uma mudança gramatical abrupta com uma gradualidade

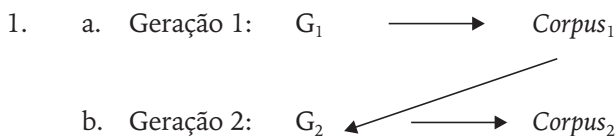
que resulta da distribuição social dos usos linguísticos. Nesse caso, a complexa distribuição social resulta da determinação de um amplo conjunto de fatores sociais sobre as escolhas que os falantes fazem entre as opções estruturais disponíveis nas gramáticas em concorrência no seio de uma mesma comunidade de fala. É claro que essa concepção se ajusta mais às situações de contato entre línguas, que estão na gênese do universo de observação das análises que integram este livro. Assim sendo, a questão da concorrência entre gramáticas no seio de uma mesma comunidade de fala constitui o objeto da próxima seção.

4.2 A competição entre gramáticas nas situações de contato entre línguas

Como já ressaltado na seção anterior, a aquisição de uma língua é um processo sempre impreciso, ou seja, nada garante que todos os aprendizes vão convergir para uma mesma gramática, mesmo a partir de um mesmo conjunto de dados primários. Segundo Kroch (2005, p.1), a *aquisição é probabilística e só aproximadamente correta*.

O modelo de aquisição da L1, em comunidades 'normais' constituídas por falantes nativos, esboçado nos trabalhos gerativistas, considera que a transmissão linguística se dá segundo o esquema a seguir:

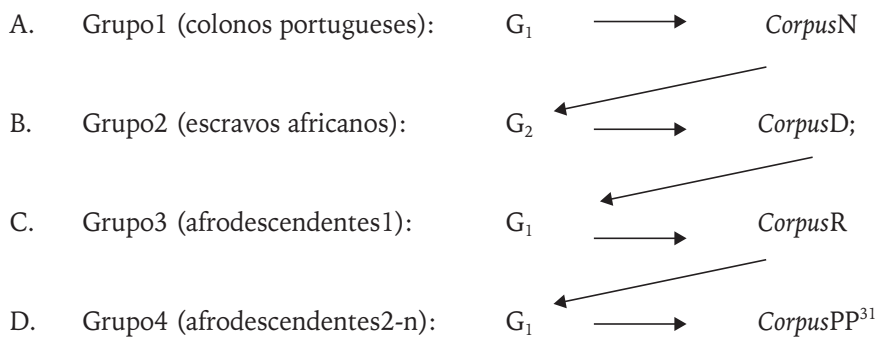
1. a. Geração 1: G_1 ⟶ *Corpus*$_1$

   b. Geração 2: G_2 ⟶ *Corpus*$_2$

A aquisição da linguagem é de natureza abdutiva, no sentido de que a criança só tem acesso aos dados de língua-E (*Corpus*$_1$ - dados do seu ambiente linguístico), mas não à G_1 (língua-I, com um conjunto de parâmetros valorados para determinadas propriedades morfossintáticas, entre outras questões). Desse modo, para qualquer população de aprendizes, podem acontecer desvios, com uma distribuição probabilística e dependente da robustez disponível para determinar propriedades gramaticais, sujeitas a variação nos dados linguísticos primários a que os aprendizes estão expostos.

Nesse contexto, as crianças cometem "erros" na aquisição da sua língua materna e demoram a corrigi-los, como observado por Kroch (2005), que faz referência a vários exemplos resultantes da pesquisa sobre aquisição, dentre os quais está a pesquisa de Clahsen (1991), que mostra que as crianças germânicas generalizam as formas nominativas para contextos de acusativos, e as formas acusativas para contextos

de dativo.[30] Como as evidências para marcação de caso são muito robustas no alemão padrão, essa generalização não apresenta indícios que levem a uma mudança na gramática adulta. Na idade entre 26 e 42 meses, as crianças se direcionam para os usos da língua-alvo. Contudo, a competição entre formas acusativas e dativas permanece até depois dos 30 meses, indicando que o processo de substituição não é instantâneo, mesmo com evidências robustas na língua-alvo.

Os fatos da história sociolinguística do Brasil (cf. capítulo 1 deste livro) revelam uma predominância de situações de multilinguismo no Brasil do século XVI ao XIX. Focalizando a questão africana, a aquisição do português como língua materna por parte dos descendentes dos escravos trazidos da África teria, em grande medida, na constituição dos PLDs, dados fragmentários e defectivos do português falado como segunda língua, em níveis muito variáveis de proficiência, o que tem sido definido aqui como um processo de **transmissão linguística irregular** (cf. capítulo 3 deste livro).

Um modelo de aquisição para situação de contato, que reflete a questão da aquisição imperfeita ou transmissão linguística irregular, é esboçado por Roberts (2007, p.390). O esquema abaixo é uma adaptação do esquema de Roberts para a situação afro-brasileira:

No esquema acima, a relação entre A e B é a de uma aquisição defectiva de segunda língua por parte de falantes adultos (os escravos africanos), tendo como *input* os dados fornecidos por falantes nativos do português (os colonos portugueses). A relação entre B e C é de aquisição da língua materna por parte dos filhos dos escravos africanos, a partir do modelo defectivo de português falado como segunda língua por estes últimos. A relação entre C e D é a de transmissão linguística geracional, que se dá entre os segmentos afrodescendentes, a partir da versão nativa de português reestruturado do Grupo 3. O *corpus*D deriva de uma

[30] Evidentemente, nem todas as crianças realizam os mesmos tipos de erros.

[31] *Corpus* PP = *corpus* do português popular.

variedade, não apenas simplificada do português falado como segunda língua (sobretudo nos mecanismos gramaticais mais abstratos e sem valor referencial), mas também com eventuais "contaminações" decorrentes de transferências, reanálises etc. orientadas pelas gramáticas das línguas nativas dos escravos africanos (cf. capítulo 3 deste livro).[32] Já o *corpus*R é constituído a partir de uma variedade nativa de português reestruturado, falado pelos escravos crioulos (filhos de mãe africana nascidos no Brasil) e refletirá, por um lado, as simplificações ocorridas na constituição do *corpus*D, bem como eventuais transferências do substrato; por outro lado, conterá também reanálises e reestruturações orientadas pelos dispositivos inatos da GU que atuam na aquisição da língua materna (cf. capítulo 3 deste livro). A partir daí, as novas gerações de afrodescendentes (o Grupo 4, dos afrodescendentes crioulos de segunda geração em diante) tenderão a ir reincorporando estruturas gramaticais do português falado pelas classes dominantes brasileiras, na medida em que as interferências do contato entre línguas diminuem (sobretudo a partir de 1850, com o fim do tráfico negreiro) e na medida em que esses indivíduos vão se integrando e ascendendo na sociedade brasileira.

Portanto, há uma relação de implicação entre os diversos níveis. Quanto maior for a simplificação/contaminação do *corpus*D, maior será a reestruturação na G1 do Grupo 3 de afrodescendentes de primeira geração. Dessa forma, o processo de transmissão da língua da segunda geração de afrodescendentes em diante se assemelhará mais a um processo de descrioulização, como o que se observa no Caribe. Porém, há que se ter em mente que, se tal processo ocorreu, ele terá sido pontual (cf. capítulo 1 deste livro). Terá predominado um processo de reestruturação leve na G1 do Grupo 3, em um grau bem menor do que o observado nas situações de crioulização típicas (cf. capítulo 3 deste livro). Uma das razões para isso é a possibilidade real de superposição entre os diversos níveis, sobretudo o maior acesso dos escravos crioulos em sua infância aos modelos dos falantes nativos do português, já que era comum que os filhos dos escravos convivessem com os filhos dos senhores, capatazes e trabalhadores livres no universo da casa-grande até cerca de sete anos de idade, quando passavam a ser empregados nas atividades produtivas (MATTOSO, 2003). Por outro lado, essa relação A-B-C foi replicada inúmeras vezes no Brasil, pois a chegada de escravos foi ininterrupta e se prolongou por mais de três séculos.

[32] É preciso ter em conta que o grau de simplificação/contaminação dessa variedade de português falado como segunda língua será muito variável em função da proficiência variável em português alcançada pelos escravos africanos no Brasil. Como destacado no capítulo 1 deste livro, os africanos empregados como escravos domésticos na casa-grande teriam muito mais proficiência em português do que os escravos empregados na lavoura e que viviam na senzala. Além disso, o nível de proficiência nas cidades seria, em média, muito maior do que no interior do país. Da mesma forma, os escravos das minas teriam, em princípio, um nível de proficiência maior do que os da lavoura, em função das especificidades de cada uma dessas atividades; e assim por diante.

4.2.1 A concorrência entre gramáticas na história sociolinguística do Brasil

Os efeitos mais visíveis da situação de interlíngua, na aquisição imperfeita por parte de falantes adultos, e de transmissão linguística irregular, em sua nativização, se refletem, sobretudo, em três importantes propriedades dos sistemas linguísticos (ROBERTS, 2007, p.389):

(i) as evidências para certos parâmetros parecem instáveis (como é o caso do parâmetro do sujeito nulo no PB culto e popular);[33]

(ii) a tendência a não realizar a morfologia flexional (como é o caso da variação de concordância no PB);

(iii) as alterações em diversas estruturas da língua-alvo (como é o caso das estruturas relativas do PB, bem como algumas partículas gramaticais, como o artigo, as preposições e os clíticos).

Dentre os aspectos linguísticos mais afetados, destaca-se a morfologia de número do Sintagma Determinante (doravante DP). Para adquirir as propriedades de marcação de número em um DP de uma língua, a criança precisa procurar pistas nos dados para decidir sobre três possibilidades lógicas, como esquematizadas abaixo, considerando apenas o DP constituído de determinante+nome:

a. D_{+PL} N_{+PL}

b. D_{+PL} N_{-PL}

c. D_{-PL} N_{+PL}

Em uma situação de aquisição de língua materna, espera-se que as três possibilidades se realizem, ao menos nas primeiras fases de aquisição, mesmo que o *input* não ofereça evidências para as três possibilidades permitidas pela GU. Em fases posteriores, a criança eliminará os usos que não são consistentes com a língua-alvo que está sendo adquirida. Em uma situação de aquisição de segunda língua, ou de contato, a variação e a permanência desses usos podem ser diferentes, a depender das situações sociolinguísticas da comunidade em que ela vive. Se uma das opções tem avaliação social positiva, a criança poderá ajustar a sua *performance* para um maior uso da variante prestigiada; se não passar por este processo de correção, porque ignora os usos de alguns falantes ou porque ouve na comunidade de fala evidências para as diferentes opções, a opção desviante pode se propagar, levando a uma mudança linguística.

Tem sido difícil testar a hipótese de transmissão linguística irregular no que diz respeito ao seu caminho diacrônico. Os testemunhos linguísticos dos

[33] Cf. capítulo 6 deste livro e Ribeiro (no prelo).

séculos anteriores praticamente se resumem a textos formais escritos por indivíduos com alto grau de letramento, cuja padronização na escrita elimina a variação da fala. Contudo, a edição de um conjunto de Atas escritas no século XIX por africanos e afrodescendentes (OLIVEIRA, 2006) revela que as três possibilidades lógicas de marcar o plural no DP estavam em variação entre os afrodescendentes no Brasil, como se pode ver nos exemplos abaixo:[34]

a. **as** forças

b. **os** conhecimento

c. **este** deveres

Tais realizações expressam uma ampla gama de flutuação que esteve na base da formação das variedades populares do português brasileiro, particularmente nas comunidades de fala afro-brasileiras, como as que são analisadas neste livro. Ainda em relação a esse amplo espectro de variação inerente às situações de multilinguismo, pode-se refletir sobre a aquisição de um item lexical com morfologia de caso, tendo por base a aquisição de caso morfológico na fala de crianças germânicas referido acima. No português trazido pelos colonos para o Brasil, o pronome *ele* possui resumidamente os seguintes traços:

a. [+pronome]

b. [-anáfora]

c. [+singular]

d. [+3ª. pessoa]

d. [+masculino]

e. [+nominativo] / [+dativo]

As alterações produzidas pela aquisição imperfeita do português como segunda língua e as abduções ocorridas na aquisição do português como língua materna nesse contexto de contato entre línguas podem estar na origem de uma série de fatos que caracterizam as variedades populares do português do Brasil atualmente:

(i) *Ele* usado com anáfora, em variedades do dialeto mineiro (LEMLE, 1985)

Ex.: João viu *ele* no espelho.

(ii) ampliação das possibilidades de caso, ocorrendo também como acusativo (DUARTE, 1989)

Ex.: Encontrei *ele* ontem.

[34] Dados do informante Gregório, afrodescendente, com 18 anos na época.

(iii) variação na distinção de gênero na relação anafórica (DETTONI, 2003, p.73-76)[35]

Ex.: A minha rua$_i$ é muito fácil de você achar *ele*$_i$, porque *ele*$_i$, bem assim na esquina, *ele*$_i$ tem uma obra da prefeitura, tem uns homens lá trabalhando.

O amplo espectro de flutuação de formas inerente aos PLDs disponíveis para a aquisição da língua materna nos contextos definidos entre os níveis A-B-C do esquema apresentado acima pode ter como consequência a concorrência de gramáticas distintas, no nível do conhecimento linguístico do falante. Nesse sentido, nas variedades de português que se vão constituindo ao longo das diversas gerações de afrodescendentes, estruturas da gramática do português das classes dominantes brasileiras convivem em concorrência com estruturas provenientes de processos de simplificações, transferências, reanálises, etc. ocorridos nas situações de contato e de transmissão linguística irregular. Dentro dessa perspectiva, a variação resultante dessa concorrência entre gramáticas deve ser vista como processos de **alternância de código** (ing. *code-switching processes*).

Na tradição da pesquisa sociolinguística, a alternância de códigos, que se manifesta tanto entre sentenças distintas (nível intersentencial), quanto no interior de uma mesma sentença (nível intrassentencial), é regulada por mecanismos discursivos e/ou pragmáticos inerentes às situações particulares de interação verbal. Já para Belazi, Ribin e Toribio (1994, p.221), a alternância de códigos no nível intersentencial seria condicionada por fatores sociais e discursivos, enquanto a alternância de códigos no interior da estrutura da sentença seria condicionada por princípios sintáticos. Assim, independentemente da dimensão discursiva do fenômeno da alternância de códigos, pode-se pensar também que ela tem implicações evidentes sobre a forma como o conhecimento linguístico do falante se estrutura no momento de sua produção verbal, sobretudo quando a alternância ocorre no interior de uma mesma sentença.

Nos casos de alternância de códigos no interior de uma mesma sentença, haverá uma tendência para predominância de uma gramática numa determinada porção da estrutura oracional e a predominância da outra gramática nos demais nódulos da oração. Abordando a questão no plano variável da fala, pode-se pensar que há uma propensão à coocorrência de estruturas provenientes de uma mesma gramática numa mesma porção da estrutura da sentença nos processos de alternância de código. Essa propensão à coocorrência pode ser definida pelo que aqui se denomina **princípio da coesão estrutural**, que será o objeto da próxima seção deste capítulo.

[35] Dettoni (2003) observa um fenômeno variável da concordância de gênero na relação anafórica no falar cuiabano, em que um antecedente feminino pode ser retomado por *ele* ou *ela*.

4.2.2 O princípio da coesão estrutural

O princípio da coesão estrutural atua fundamentalmente nos casos de alternância de código determinados pela concorrência de duas gramáticas presentes na mente/cérebro do falante, que se atualiza no seu desempenho linguístico na forma de uma **variação estruturada**. Em muitos casos de alternância de código, estruturas geradas a partir de regras de uma gramática podem-se combinar com estruturas geradas segundo princípios da outra gramática, no interior de uma mesma sentença. As restrições e os limites para a ocorrência de alternância de código na projeção estrutural de uma sentença levantam, assim, questões relevantes para a teoria da gramática.

Há uma série de estudos que apontam para restrições à ocorrência de alternância de códigos nos mais variados níveis da sentença. Assim, por exemplo, Gumperz (1970 *apud* POPLACK, 1980b) e Timm (1975 *apud* POPLACK, 1980b) referem-se a restrições à alternância de códigos entre um sujeito pronominal e o verbo, entre verbos auxiliares e verbos plenos, entre verbos e objetos, e entre elementos negativos e verbos. Num nível mais alto da representação estrutural da sentença, admitem-se ainda restrições apontadas entre complementizadores e sentenças encaixadas (GUMPERZ, 1976 apud POPLACK, 1980b) e entre certos elementos interrogativos e verbos (WOOLFORD, 1983). Numa visão ainda mais restritiva, Ribeiro (1996, p.329) afirma que "a alternância só será possível nos pontos das representações estruturais abstratas em que não haja violação dos seus princípios licenciadores, para todo e qualquer par de línguas"; e justifica a sua posição da seguinte maneira:

> Se admitimos ser a Teoria dos Princípios e Parâmetros um modelo de gramática ótimo para as línguas humanas, temos de reconhecer que também o é para os enunciados com alternância de código, visto serem eles resultantes da justaposição de duas gramáticas num enunciado. Concebemos, desse modo, que as alternâncias de códigos só devem ocorrer nos limites sintáticos em que a justaposição de L1 e L2 não viole nenhum dos princípios dos diversos subsistemas da gramática, ou seja, a alternância de código deve ser licenciada por princípios da UG (RIBEIRO, 1996, p.358).

Dentro dessa perspectiva, o *principio de coesão estrutural* estaria relacionado com restrições à ocorrência de alternância de código no interior da estrutura da sentença, mais fortes nos nódulos nucleares do que nos nódulos periféricos.

Por outro lado, o fenômeno da alternância de códigos seria condicionado também por uma série de fatores relativos à situação em que se dá a interação verbal e às características dos indivíduos nela envolvidos. Poplack (1980b) destaca a importância do nível de bilinguismo dos falantes. Para ela, a alternância de

códigos no interior da sentença ocorreria normalmente quando os falantes apresentassem um alto grau de proficiência nas duas línguas; enquanto a alternância intersentencial não demandaria o mesmo grau de competência. Belazi (1991) demonstra também que os condicionamentos gramaticais da alternância de códigos seriam mais sensíveis à medida que aumenta o grau de competência linguística dos falantes em ambas as línguas.

Um outro fator que condicionaria o processo de alternância de códigos no interior da sentença, visto assim como um fenômeno variável, seria o grau de proximidade das gramáticas envolvidas. Em se tratando de variedades linguísticas muito distintas em termos tipológicos, as restrições e os limites tendem a ser mais nítidos, de modo que a alternância tende a ser mais estrita. Já nas situações em que estão em contato, não duas línguas distintas, mas duas variedades da mesma língua histórica, as fronteiras que se podem estabelecer para a alternância de códigos são bem mais tênues e muito menos nítidas.

Essa seria a situação que se verifica, por exemplo, no caso das comunidades rurais afro-brasileiras focalizadas neste livro, em que se observa a convivência de duas gramáticas muito semelhantes e intimamente relacionadas. A rigor, estruturas de uma gramática anterior (mais afetadas pelas alterações decorrentes das situações de contato entre línguas) se manteriam sob a concorrência, em crescente predomínio, de estruturas da gramática das variedades urbanas do português brasileiro, no interior de um mesmo sistema heterogêneo, de modo que a forma como as estruturas geradas pelas duas gramáticas se distribuem no interior da sentença está muito longe de ser regida por princípios rígidos e categóricos.

Considerando-se este contexto de estruturas subjacentes tão aparentadas em que a situação de concorrência entre duas gramáticas chega a ser dúbia e objeto de uma larga controvérsia teórica,[36] o *princípio da coesão estrutural* pode ser formulado com base na propensão de coocorrência de estruturas de uma mesma gramática numa determinada porção da sentença, nas situações de alternância de

[36] Nesse ponto, abre-se um extenso debate teórico sobre a configuração da competência linguística do falante individual, em que se opõem as concepções da Sociolinguística Variacionista, de um lado, e o programa gerativista, de outro. As formulações dessa última vertente (como, por exemplo, Kroch e Taylor, 1997), ao assumirem a concorrência de duas gramáticas, assumem um sistema formal homogêneo, no qual a variação é unicamente social, derivada de situações de diglossia e de bilinguismo. Já a Sociolinguística (cf. WLH, 1968 [1989]), ao assumir um sistema heterogêneo, define uma gramática com regras opcionais, na qual as opções feitas pelos falantes é que são condicionadas por fatores sociais. Este impasse teórico transcende os limites deste capítulo, que, como foi dito em seu início, não pretende esgotar uma questão teórica de tal envergadura. Entretanto, um ponto importante deve ser destacado: os limites da opcionalidade formal da gramática não se estende para qualquer fenômeno linguístico, fora das situações de multilinguismo que reúnem línguas tipologicamente muito diferenciadas. E, como também já foi dito acima, o escrutínio dos limites da variação e da invariância na gramática das diversas comunidades de fala pode-se tornar o campo mais profícuo da pesquisa linguística na atualidade.

código, e não numa coocorrência categórica, estritamente restringida pelos princípios abstratos da gramática. Essa propensão seria condicionada, em primeiro lugar, pelo grau de coesão da estrutura delimitada. Assim, como já foi dito, haveria uma maior propensão de coocorrência de estruturas de uma mesma gramática em seus nódulos mais coesos do que nos nódulos mais periféricos.[37]

De qualquer forma, é possível fazer previsões probabilísticas, sobretudo em níveis estruturais mais coesos, como o da estrutura sintagmática de um mesmo DP. Nesse caso, pode-se pensar, tomando-se como exemplo as regras de concordância nominal, que a regra de concordância de gênero tende a ser menos aplicada em um DP em que não se aplica a regra de concordância de número, do que em um DP em que essa regra é aplicada, de modo que, no universo linguístico das comunidades rurais afro-brasileiras isoladas, um DP como *as coisas mais caro* constitui uma ocorrência bem factível, ao passo que um DP do tipo *as coisa mais caros* é muito improvável. Isso decorre do fato de que a concordância de número é mais afetada no processo de transmissão linguística irregular do que a concordância de gênero (cf. capítulos 11 e 14 deste livro). Assim, é possível postular uma relação de implicação. Ou seja, se em um DP o falante chega a aplicar a regra de concordância de número plenamente, a regra de concordância de gênero será aplicada num nível categórico; contudo, o inverso não seria verdadeiro, podendo o falante aplicar a regra de concordância de gênero em um DP, sem que isso leve à aplicação da regra de concordância de número, de modo que um DP como *as coisa mais cara* é muito frequente.

Já no nível das relações entre constituintes no esqueleto estrutural da sentença, o princípio da coesão estrutural também se aplica, porém com um grau de implicação mais diluído. Tal é o caso da correlação entre a aplicação da regra de concordância de número no DP sujeito e a aplicação da regra de concordância de pessoa e número entre o verbo e este DP sujeito, que especifica os traços de pessoa e número do verbo (cf. capítulo 12 deste livro). Dessa forma, o princípio da coesão estrutural prevê que o falante tenderá a fazer mais a concordância verbal quando fizer a concordância nominal. Assim, um SN como *o**s** menino**s*** atuará como um fator mais favorecedor da aplicação da concordância verbal do que um SN como *o**s** menino*. Desse modo, essa correlação se atualiza de forma positiva em frases como (1) e (2):

(1) O**s** menino**s** já *cheg**aram***.

(2) Os menino já cheg**ou**.

[37] Nódulos coesos significam aqui os que estão internos a um único constituinte, como os diferentes nódulos de manifestação de traços-phi de um DP (número, gênero, pessoa); nódulos periféricos referem-se não só ao tradicional conceito de adjunção, mas, sobretudo, à relação que se estabelece entre um constituinte e o núcleo de um outro nódulo, como, por exemplo, à relação entre um DP sujeito e o núcleo da flexão sentencial (TP), ambos estando em uma relação de especificador / núcleo.

Entretanto, nesse nível da estrutura, o efeito do princípio da coesão estrutural é mais tênue do que no nódulo mais coeso da estrutura interna do DP, de maneira que a correlação prevista pode não ocorrer. Mas, mesmo nesse caso, predições podem ser feitas, em termos implicacionais. É mais provável que o falante aplique a regra de concordância verbal sem aplicar a regra de concordância nominal — cf. exemplo (3) — do que o contrário — cf. exemplo (4). Tal propensão é em grande parte determinada pelo fato de o grau de variação no uso da regra de concordância nominal no português brasileiro ser maior do que o grau de variação no uso da regra de concordância verbal.

(3) O**s** menino já *cheg**aram***.

(4) O**s** menino**s** já *chegou*.

Tendo como fundamento esta concepção da variação linguística dentro da estrutura gramatical, o **princípio da coesão estrutural** será empregado em várias análises que constam da segunda parte deste livro. Nessa aplicação, tal princípio não será tomado em um sentido teórico mais forte, como um fator rígido e categórico, mas em um sentido mais fraco, isto é, probabilístico. Os efeitos probabilísticos desse princípio, bem como de outros fatores linguísticos, foram mensurados, na maioria das análises deste livro, com base no suporte metodológico comumente usado nas análises variacionistas. Esse e os demais procedimentos metodológicos que caracterizam o conjunto das pesquisas que deram ensejo a este livro serão descritos no próximo capítulo.

5 A metodologia

Dante Lucchesi

Para que se possa analisar a variação nos padrões coletivos de comportamento linguístico no interior de uma comunidade de fala de forma satisfatória, a investigação deve atender a certas exigências metodológicas. Em primeiro lugar, é preciso quantificar as distintas possibilidades de expressão (que definem a variação linguística), em função dos contextos estruturais e das situações de uso da língua (o que define a sua sistematicidade). Portanto, quanto mais refinada for a quantificação das ocorrências analisadas, melhores serão as condições para que se possa explicar o fenômeno variável. Grande parte dos estudos que compõem a segunda parte deste livro utilizou, como suporte para a análise quantitativa dos dados, o **Pacote de Programas VARBRUL**, que será brevemente apresentado na seção 5.2 deste capítulo.

Por outro lado, essa base de dados deve ser extraída de uma amostra de fala que reflita, tanto quanto possível, as condições mais naturais de uso linguístico, pois o objeto da investigação deve ser situado no processo histórico real de constituição e atualização da língua. Ou seja, a investigação deve se basear em uma amostra do que Labov (1972b) definiu como o *vernáculo* da comunidade. Além disso, as dimensões da amostra devem ser suficientes para que se possa observar, com uma boa margem de segurança, a variação no tópico analisado, bem como para que possam ser consideradas as influências de certos fatores sociais, tais como: a idade, o sexo e o nível de escolaridade do falante. As técnicas empregadas na constituição da amostra de fala em que se baseiam as análises contidas neste livro, bem como as suas dimensões, serão apresentadas na seção que se segue.

5.1 O acervo de fala vernácula do português afro-brasileiro

Na constituição do **Acervo de Fala Vernácula do Português Afro-Brasileiro do Estado da Bahia**, foram definidas quatro comunidades rurais afro-brasileiras isoladas onde seriam recolhidas as amostras de fala vernácula: as comunidades geminadas de Barra e Bananal, no Município de Rio de Contas;

Helvécia, no Município de Nova Viçosa; Cinzento, no Município de Planalto; e Sapé, no Município de Planalto.[1] Em cada comunidade, os pesquisadores do **Projeto *Vertentes do Português Popular do Estado da Bahia*** (doravante Projeto *Vertentes*)[2] realizaram entrevistas com pelo menos doze de seus moradores escolhidos aleatoriamente, entre os que nasceram na comunidade. Nessas entrevistas, o pesquisador procurava desenvolver uma conversação a mais informal possível com o membro da comunidade por um período de no mínimo 40 minutos e no máximo uma hora (cf. seção 5.1.2, abaixo).

A escolha dos informantes em cada comunidade de fala atendeu às seguintes variáveis estratificadas: sexo (masculino/feminino) e idade (faixa 1: de 20 a 40 anos; faixa 2: de 40 a 60 anos; faixa 3: de mais de 60 anos). A amostra constituía-se, assim, de seis células, com dois informantes em cada célula, num total de 12 informantes por comunidade assim distribuídos: duas mulheres e dois homens de 20 a 40 anos; duas mulheres e dois homens de 40 a 60 anos; duas mulheres e dois homens de mais de 60 anos; e um total de 48 entrevistas no conjunto das quatro comunidades selecionadas.

As entrevistas das comunidades de Rio de Contas, Helvécia e Cinzento foram gravadas em meio magnético (fitas cassete) e depois digitalizadas. As entrevistas de Sapé já foram feitas diretamente em meio digital (gravadores de *minidisk*). Todas as entrevistas foram editadas em meio digital para melhorar a qualidade do áudio e eliminar ruídos (cf. seção 5.1.4). Depois foram transcritas grafematicamente de acordo com os critérios definidos no Projeto (cf. seção 5.1.5). As transcrições foram revistas utilizando-se os meios informáticos de tratamento de áudio, em ambiente *Windows*.

Desse conjunto de 48 entrevistas armazenadas em meio digital e suas transcrições, que constitui o **Acervo de Fala Vernácula do Português Afro-Brasileiro do Estado da Bahia**, foram selecionadas 24 entrevistas que passaram a constituir o ***Corpus* Base do Português Afro-Brasileiro do Estado da Bahia**. Para tanto, foram escolhidos dois informantes de cada uma das três faixas etárias (sendo um de cada sexo) de cada uma das quatro comunidades integrantes do *Acervo*, de modo a equilibrar a diferença geracional na amostra, guardando-se um intervalo médio de aproximadamente 20 anos entre uma faixa etária e outra (cf. Quadros 1 e 2). A esse conjunto de 24 entrevistas foram acrescentadas outras quatro, com os dois informantes mais velhos de cada sexo. Essa faixa 4, constituída por falantes com mais de 80 anos, permite visualizar os estágios mais antigos da gramática das

[1] A descrição dessas comunidades é feita no capítulo 2 deste livro.

[2] Uma descrição detalhada do Projeto *Vertentes* pode ser encontrada na Internet, no seguinte endereço: <http://www.vertentes.ufba.br>.

comunidades de fala analisadas,[3] com o intuito de se identificarem os traços mais ligados aos efeitos do contato entre línguas que marca a origem dessas comunidades.

O estabelecimento desse *corpus* visou, por um lado, a viabilizar uma base de dados mais factível para as análises linguísticas de aspectos muito recorrentes na fala. Por outro lado, buscou-se também estabelecer uma distribuição mais equilibrada dos informantes segundo as variáveis escolaridade e estada fora da comunidade. Essas duas outras variáveis sociais foram definidas nos seguintes termos. Na variável escolaridade, fez-se a distinção entre falantes analfabetos e semianalfabetos, sendo estes últimos aqueles que tiveram qualquer experiência com o universo do letramento, mesmo que o resultado disso se resumisse à capacidade de *assinar o nome*. Já na variável estada fora da comunidade, fez-se a distinção entre os indivíduos que já tinham vivido pelo menos seis meses fora da comunidade (geralmente por motivo de trabalho), e aqueles que sempre moraram na comunidade. Entretanto, não se conseguiu uma distribuição perfeita, em que todas as células fossem preenchidas. Não se dispunha, por exemplo, de mulheres semialfabetizadas com mais de sessenta anos. Já entre os homens jovens, sucedeu o contrário: era difícil encontrar analfabetos completos. Tais lacunas do *corpus* constituído nada mais são do que o reflexo do universo sociocultural em que foi feita a recolha.

5.1.1 A estrutura das amostras de fala vernácula

Nos Quadros 1 a 3, estão apresentadas as características do ***Corpus* do Português Afro-Brasileiro do Estado da Bahia**, juntamente com as características das amostras recolhidas nas quatro comunidades, e cada entrevista é indicada da seguinte maneira:

(1) **Comunidade**: **RC**, Rio de Contas; **HV**, Helvécia; **CZ**, Cinzento; **SP**, Sapé.

(2) **Número da entrevista, em sua amostra original**: **RC-Inq04**, por exemplo.

(3) **Sexo do informante**: **F**, feminino; **M**, masculino.

(4) **Idade do informante**: **26a**, vinte e seis anos, por exemplo.

(5) **Nível de escolaridade do informante**: **A**, analfabeto; **S**, semianalfabeto.

(6) **Estada fora da comunidade**: **E**, o informante viveu pelo menos seis meses fora da comunidade; **N**, o informante viveu sempre na comunidade.

[3] Com base na hipótese clássica da abordagem em **tempo aparente** (NARO, 2003b, p. 44), "o estado atual da língua de um falante adulto reflete o estado da língua adquirida quando o falante tinha aproximadamente 15 anos de idade". Não obstante a possibilidade de levantar alguns questionamentos a essa correlação (LUCCHESI, 2001b), pode-se pensar que, a grosso modo, a gramática de um falante com 85 anos de idade, na década de 1990, remete para a situação da gramática da comunidade de fala em torno da década de 1920.

FAIXA 1					FAIXA 2					FAIXA 3				
RC-Inq04	F	37a	A	E	RC-Inq13	F	47a	S	N	RC-Inq24	F	75a	A	N
HV-Inq01	F	29a	A	N	HV-Inq07	F	42a	A	E	NHV-Inq13	F	85a	A	N
SP-Inq01	F	24a	S	N	SP-Inq05	F	53a	S	E	NSP-Inq09	F	76a	A	E
CZ-Inq01	F	28a	S	E	CZ-Inq08	F	50a	A	N	CZ-Inq09	F	63a	A	N
(média: 29,5 anos)					(média: 45,5 anos)					(média: 74,75 anos)				

Quadro 1 - Projeto *Vertentes* - *Corpus* base do português rural afro-brasileiro: MULHERES

FAIXA 1					FAIXA 2					FAIXA 3				
RC-Inq05	M	26a	S	E	RC-Inq08	M	55a	A	E	RC-Inq26	M	68a	S	E
HV-Inq04	M	30a	S	N	HV-Inq12	M	57a	A	N	HV-Inq20	M	70a	A	E
SP-Inq04	M	28a	A	N	SP-Inq06	M	42a	S	N	SP-Inq12	M	66a	A	N
CZ-Inq03	M	34a	S	E	CZ-Inq06	M	48a	S	E	CZ-Inq11	M	64a	S	N
(média: 29,5 anos)					(média: 50,5 anos)					(média: 67 anos)				

Quadro 2 - Projeto *Vertentes* - *Corpus* base do português rural afro-brasileiro: HOMENS

FAIXA 4									
HV-Inq19	F	103a	A	E	HV-Inq22	M	80a	A	N
CZ-Inq10	F	107a	A	N	CZ-Inq12	M	82a	A	N
(média: 93 anos - 22,125 anos de diferença sobre a Faixa 3)									

Quadro 3 - Projeto *Vertentes* - *Corpus* base do português rural afro-brasileiro: MULHERES e HOMENS com mais de 80 anos

Observações (assimetrias não superadas no *corpus*):

1) Há três homens semianalfabetos na faixa 1
2) Todas as mulheres da faixa 3 são analfabetas.

Média de idade por faixa etária:

Faixa 1 29,5 anos

Faixa 2 49,25 anos (19,75 anos de diferença sobre a Faixa 1)

Faixa 3 70,875 anos (21,625 anos de diferença sobre a Faixa 2)

FAIXA 1					FAIXA 2					FAIXA 3				
RC-Inq01	F	26a	S	E	RC-Inq08	M	55a	A	E	RC-Inq20	M	77a	A	N
RC-Inq04	F	37a	A	E	RC-Inq09	M	43a	S	E	RC-Inq21	F	74a	A	N
RC-Inq05	M	26a	S	E	RC-Inq11	F	42a	S	E	RC-Inq24	F	75a	A	N
RC-Inq06	M	36a	S	E	RC-Inq13	F	47a	S	N	RC-Inq26	M	68a	S	E
RC-Inq07	F	26a	S	N										

Quadro 4 - Projeto *Vertentes* - *Corpus* base do português rural afro-brasileiro: RIO DE CONTAS

FAIXA 1					FAIXA 2					FAIXA 3				
HV-Inq01	F	29a	A	N	HV-Inq07	F	42a	A	E	HV-Inq13	F	85a	A	N
HV-Inq02	M	28a	S	E	HV-Inq08	M	44a	S	N	HV-Inq15	F	80a	A	N
HV-Inq03	F	35a	S	E	HV-Inq09	F	54a	A	N	HV-Inq19	F	103a	A	E
HV-Inq04	M	30a	S	N	HV-Inq10	M	50a	S	N	HV-Inq20	M	70a	A	E
HV-Inq05	F	38a	S	E	HV-Inq11	F	60a	A	N	HV-Inq22	M	80a	A	N
HV-Inq06	M	39a	S	N	HV-Inq12	M	57a	A	N	HV-Inq24	M	67a	A	N

Quadro 5 - Projeto *Vertentes - Corpus* base do português rural afro-brasileiro: HELVÉCIA

FAIXA 1					FAIXA 2					FAIXA 3				
SP-Inq01	F	24a	S	N	SP-Inq05	F	53a	S	E	SP-Inq09	F	76a	A	E
SP-Inq02	M	27a	S	N	SP-Inq06	M	42a	S	N	SP-Inq10	M	70a	A	N
SP-Inq03	F	36a	A	N	SP-Inq07	F	55a	A	N	SP-Inq11	F	77a	A	N
SP-Inq04	M	28a	A	N	SP-Inq08	M	48a	S	E	SP-Inq12	M	66a	A	N

Quadro 6 - Projeto *Vertentes - Corpus* base do português rural afro-brasileiro: SAPÉ

FAIXA 1					FAIXA 2					FAIXA 3				
CZ-Inq01	F	28a	S	E	CZ-Inq05	M	41a	S	N	CZ-Inq09	F	63a	A	N
CZ-Inq02	F	16a	S	N	CZ-Inq06	M	48a	S	E	CZ-Inq10	F	107a	A	N
CZ-Inq03	M	34a	S	E	CZ-Inq07	F	44a	S	N	CZ-Inq11	M	64a	S	N
CZ-Inq04	M	41a	S	E	CZ-Inq08	F	50a	A	N	CZ-Inq12	M	82a	A	N

Quadro 7- Projeto *Vertentes - Corpus* base do português rural afro-brasileiro: CINZENTO

Observação: As amostras de Rio de Contas e Helvécia contam com mais de doze entrevistas (treze e dezoito, respectivamente).

Esse acervo de fala forneceu a base empírica para o conjunto de análises contidas neste livro, porém cada análise definiu a(s) comunidade(s) que observaria e a dimensão de sua amostra de fala vernácula, de acordo com os seus objetivos específicos.

5.1.2 O caráter das entrevistas

A recolha das amostras de fala em cada comunidade foi realizada por meio de entrevistas conduzidas por pesquisadores do Projeto *Vertentes* junto aos seus moradores, buscando superar o **paradoxo do observador**, já que o objetivo era extrair o **vernáculo** da comunidade (LABOV, 1966, 1972b). Para isso, buscou-se criar um clima de descontração durante a realização das entrevistas, que, no geral,

foram realizadas na casa dos informantes, ou "na roça" (local de cultivo situado na pequena propriedade do informante ou na propriedade de outrem); e foram conduzidas pelos pesquisadores, contando muitas vezes com o apoio de um membro da comunidade, contratado previamente para esse fim. Os temas levantados pelos inquiridores foram aqueles que provocavam um maior envolvimento do informante com o que estava sendo falado, desviando-o da forma como estava sendo falado, em conformidade com os princípios metodológicos da pesquisa sociolinguística.

5.1.3 Os temas das entrevistas

Labov (1972a) demonstrou que a narrativa de experiências pessoais é o tópico mais produtivo para a conversação espontânea. Ao narrar as suas experiências de vida, o falante naturalmente se envolve com o referente da conversa, desligando-se da forma como verbaliza esse referente. Desse modo, quanto mais envolvente for o tipo de experiência narrada, quanto mais de perto ela tocar o falante, melhor será o nível da entrevista. Seguindo esse raciocínio, Labov e seus assistentes desenvolveram módulos de perguntas sobre temas dessa natureza, tendo o módulo sobre o tema *risco de vida* se revelado um dos mais produtivos.

Na constituição do acervo de fala do português afro-brasileiro, o tema *risco de vida* foi adaptado para a realidade das comunidades pesquisadas através do tema *animais peçonhentos*. São raros os informantes adultos que não tenham alguma experiência com picaduras de animais como cobras, lacraias e escorpiões, de modo que a abordagem do tema quase sempre suscitava uma boa narrativa de experiência pessoal seguida de uma detalhada exposição sobre os recursos tradicionalmente usados na comunidade contra as picaduras.

Não obstante a produtividade de alguns temas, o leque de temas numa entrevista de tipo sociolinguístico é potencialmente ilimitado, não havendo em princípio tema ruim. O fator decisivo é a sensibilidade do pesquisador em capturar o estado psicológico do informante, tendo como apoio uma boa avaliação da realidade cultural da comunidade. Abaixo são listados os temas mais invocados no trabalho de constituição de amostras de fala vernácula nas comunidades estudadas:

(i) História da comunidade: "a vida antigamente", "os costumes dos antigos", etc.

(ii) A infância: "as experiências da infância", "as brincadeiras", "a relação com os pais", etc.

(iii) Doenças: tema particularmente produtivo, pois foram obtidas boas narrativas sobre o acompanhamento de parentes adoentados, em função da dificuldade do acesso ao atendimento médico nessas comunidades.

(iv) Morte na família: tema igualmente produtivo.

(v) Emigração: geralmente o informante tinha parentes que tinham ido tentar a vida fora.

(vi) Festas locais: a descrição dessas festas, bem com a participação do informante nelas também renderam boas narrativas.

(vii) Elementos da cultura musical: em Rio de Contas, por exemplo, pôde-se obter algumas descrições do *samba* e do *bendengó*; já em Cinzento, foram obtidas boas narrativas do *reisado*.

(viii) Festas religiosas: "novena de Natal", "dias santos", "Páscoa", etc.

(ix) Intervenções exteriores: nos arraiais de Rio de Contas, todos os informantes tinham alguma coisa para falar sobre a barragem construída na região, e o modo como ela havia afetado as suas vidas; em Helvécia, era o impacto ambiental da implantação das grandes plantações de eucalipto na região.

(x) Relação com comunidades e centros vizinhos: relações pessoais, de parentesco, comerciais (idas à feira), etc.

(xi) Casamentos, namoros, vida conjugal.

(xii) Experiência de vida fora da comunidade: com base na experiência de muitos informantes (principalmente os homens) que viveram fora, trabalhando em grandes centros urbanos.

(xiii) A lavoura: "as melhores lavouras", "as chuvas e a seca", "as técnicas de cultivo, de produção artesanal da farinha de mandioca, do melaço da cana-de-açúcar e de outros produtos", etc.

(xiv) Narrativas populares: contos e histórias da tradição oral têm sido bastante usados na crioulística e constituem um bom material para a análise de estruturas mais arcaizantes, dado o seu caráter naturalmente conservador, mas foi difícil obter esse tipo de narrativa.

Como foi dito acima, a definição do leque de temas em cada recolha tinha por base o máximo de informações que se conseguia obter sobre a história e a cultura da comunidade estudada. Além disso, era preciso estar atento, no curso da interação, para os temas mais produtivos e para aqueles em que os membros da comunidade assumiam uma atitude mais refratária. O tema da escravidão, por exemplo, que se mostrou bastante produtivo em Helvécia, resultou nulo em Rio de Contas. Isso se deve a um provável pacto de silêncio que se firmou entre escravos foragidos e garimpeiros clandestinos quando da ocupação da área dos arraiais.

Mesmo dentro de uma única comunidade, o mesmo tema podia ser produtivo com uns informantes e improdutivo com outros, devendo o pesquisador estar sensível a cada situação. O tema de "entidades do sobrenatural", por exemplo, provocava uma reação negativa entre os falantes mais velhos dos arraiais de Rio de

Contas, ao passo que era bastante motivador com alguns dos falantes mais jovens. Em última instância, o processo acabou por ser bastante empírico.

5.1.4 O tratamento do áudio

Como todas as entrevistas do Projeto *Vertentes* foram realizadas de modo informal, os locais em que elas se realizaram foram os mais variados, incluindo plantações e pastos, casa de farinha, casa do informante ou seu quintal; em suma, qualquer lugar onde se pudesse desenvolver uma conversação informal. Com isso, muitas vezes a gravação da entrevista está entremeada de ruídos indesejados (som de animais, de criança chorando, do vento, etc.). A utilização de recursos informatizados para a digitalização do áudio da entrevista permitiu eliminar alguns desses ruídos, tornando a entrevista mais inteligível.

As entrevistas nas três primeiras comunidades rurais afro-brasileiras isoladas estudas pelo Projeto *Vertentes* (Rio de Contas, Helvécia e Cinzento) foram originalmente gravadas, e continuaram armazenadas, em fitas cassete, um meio analógico. Mas todo esse material foi também digitalizado. Desse modo, os acervos de fala vernácula do Projeto *Vertentes* estão sendo armazenados em meio digital, conservando-se de forma segura e duradoura um importante acervo para a pesquisa da realidade linguística do país.

5.1.5 A transcrição

O levantamento das ocorrências para a constituição da base de dados das análises não foi feito diretamente através da audição das entrevistas, mas a partir do texto da sua transcrição grafemática. Nessa transcrição, buscou-se ao máximo preservar as características da fala dos informantes; notadamente no nível da morfossintaxe. Assim sendo, as entrevistas foram transcritas grafematicamente de acordo com os critérios da **Chave de Transcrição do Projeto *Vertentes***.[4] A transcrição bem como o áudio digitalizado das entrevistas estão armazenados na sala do Projeto *Vertentes* no Instituto de Letras da Universidade Federal da Bahia. Os dados extraídos desse material foram processados quantitativamente com o recurso ao **Programa das Regras Variáveis (VARBRUL)**, na grande maioria das análises contidas neste livro.

[4] A Chave de Transcrição do Projeto Vertentes está acessível na Internet em: <http://www.vertentes.ufba.br/chave.htm>.

5.2 O suporte quantitativo

A visão teórica de que os padrões de comportamento linguístico são o produto da ação dos diversos fatores intervenientes na atividade linguística concreta apoia-se na análise empírica das probabilidades associadas aos usos linguísticos. A correlação dos fatores que atuam sobre a atividade linguística pode ser disposta em dois planos: de um lado, estão os fatores associados às características sociais do falante (tais como idade, sexo, nível de escolaridade, etc.); de outro lado, estão os fatores que constituem os contextos linguísticos nos quais o falante opera a escolha entre duas ou mais formas linguísticas para expressar a mesma informação. As formas diferentes que expressam o mesmo conteúdo semântico são denominadas **variantes**. O conjunto dessas variantes constitui a **variável linguística** a ser analisada. Portanto, a escolha subjacente a cada ocorrência da variável analisada é condicionada por uma série de fatores da estrutura linguística e da estrutura social da comunidade de fala.

Entretanto, a ação de cada um desses fatores não é isolada, ou seja, não se dá de forma independente da ação dos demais fatores, já que cada contexto de ocorrência da variável analisada resulta da combinação específica dos valores igualmente variáveis de cada um dos fatores intervenientes. O problema, então, está em isolar a interferência de cada fator independentemente dos demais, vez que, nos contextos reais, a ação de todos os fatores é simultânea (CEDERGREN; SANKOFF, 1974).

Em função disso, no âmbito da pesquisa sociolinguística, uma série de modelos matemáticos foram desenvolvidos para formalizar analiticamente a atuação combinada dos diversos fatores integrados na análise — de modo a se poder aferir a atuação específica de cada fator —, até se chegar à formulação do **modelo misto ou logístico**, proposto por Rousseau e Sankoff (1978) e "considerado mais adequado para a análise de fenômenos variáveis" (SCHERRE, 1996).[5] Esse modelo logístico, contido no pacote de programas **VARBRUL** (do inglês *variable rules*), criado por David Sankoff (SANKOFF, 1988; PINTZUK, 1988), foi utilizado por grande parte das análises contidas na segunda parte deste livro no processamento quantitativo dos dados, que visa a aferir a atuação específica dos diversos fatores linguísticos e extralinguísticos em cada fenômeno analisado.[6]

A análise quantitativa realizada pelo VARBRUL parte da codificação de cada ocorrência da variável linguística analisada (dita **dependente**) com base nos valores atribuídos aos fatores linguísticos e sociais (denominados **variáveis independentes**) previamente selecionados pelo pesquisador e produz:

---

[5] Naro (2003a) faz uma clara exposição sobre o desenvolvimento do modelo logístico.

[6] Uma análise mais detalhada do funcionamento do VARBRUL pode ser encontrada em Scherre e Naro (2003).

(i) as frequências de uso associadas a cada um dos valores das variáveis independentes;

(ii) a seleção das variáveis independentes estatisticamente relevantes;

(iii) o peso relativo de cada um dos valores das variáveis independentes com valor estatístico;

(iv) o nível de significância dos resultados obtidos, com base no **logaritmo da verossimilhança**.

As frequências referidas em (i) são o primeiro resultado fornecido pelo VARBRUL e representam as porcentagens brutas da frequência de uso de cada variante da variável analisada associadas a cada um dos níveis dos fatores arrolados previamente. A partir daí, a análise quantitativa prossegue por níveis sucessivos (num processo denominado *step up*). Cada nível é definido pela escolha de uma variável independente, cujos resultados serão combinados com os resultados obtidos nas demais variáveis, até que não reste nenhuma variável estatisticamente relevante. A seleção das variáveis estatisticamente relevantes é feita com base em seu **nível de significância**, ou seja, o índice de adequação dos resultados estatísticos projetados aos resultados efetivamente observados, gerado por um parâmetro estatístico denominado *log likelihood*, ou o **cálculo da verossimilhança máxima** (SCHERRE; NARO, 2003, p.165 et seq.).

O produto final é constituído pelo *peso relativo* de cada um dos valores de todas as variáveis independentes selecionadas pelo programa como estatisticamente relevantes. Os pesos relativos medem, numa escala de zero a 99, a interferência daquele fator sobre a variável analisada, com base numa equação que consegue integrar a influência simultânea dos demais fatores. O princípio geral é o de que os valores acima de .50 indicam uma ação favorecedora à realização da variante em foco; os valores inferiores, uma ação desfavorecedora; e os valores próximos a .50 apontariam para uma neutralidade do fator. Porém, uma leitura mais acurada deve tomar o peso relativo de cada fator relativamente aos pesos dos demais fatores do mesmo grupo.

Os resultados finais também devem estar dentro da margem de segurança definida pelo **nível de significância** (o que lhes confere confiabilidade estatística). Ou seja, o nível de segurança dos resultados finais, bem como dos resultados de cada variável independente no momento da sua seleção, deve ser igual ou inferior a .050.

Não obstante a valiosa contribuição de um programa como esse para mensurar os fenômenos analisados, não se podem tomar os resultados quantitativos como um retrato fiel da realidade. Metaforicamente, pode-se pensar nesses resultados como uma foto da realidade, e não como os fatos em si. Desse modo, o produto final da análise deve resultar da mediação da interpretação crítica do analista sobre os resultados brutos. Esse foi o espírito que presidiu a grande maioria das análises que se apresentam na segunda parte deste livro.

2ª Parte

A gramática do português afro-brasileiro

6 A realização do sujeito pronominal

*Dante Lucchesi**

A reestruturação gramatical provocada pelas situações de contato maciço entre línguas atua sobre a estrutura da variedade linguística que se forma em tais situações de três formas: (i) na redução da morfologia flexional; (ii) na alteração da marcação de parâmetros sintáticos em função de parâmetros não marcados; e (iii) na gramaticalização de itens lexicais para substituir os elementos gramaticais perdidos no estágio inicial do contato (cf. capítulo 3). A redução na morfologia flexional, sobretudo a morfologia relacionada às regras de concordância, é o processo mais geral dos três, sendo o mais notável no português brasileiro (cf. capítulos 11 e 14). Por outro lado, a gramaticalização, referida em (iii), que, na sua amplitude e intensidade, pode ser apontada como a principal característica das línguas crioulas *vis-à-vis* à estrutura de suas línguas lexificadoras, está virtualmente ausente do português afro-brasileiro, só se manifestando lateralmente. Por fim, os processos de alteração nos valores dos parâmetros sintáticos constituem a parte mais nebulosa da questão de como as situações de contato afetam a estrutura gramatical da língua (cf. capítulo 3).

O panorama sociolinguístico do Brasil fornece um quadro muito nítido em que um maior nível de contato na formação de uma variedade linguística determina uma menor frequência de uso da morfologia flexional às regras de concordância, como se pode constatar na amplitude estrutural do processo de variação observado nas comunidades rurais afro-brasileiras isoladas *vis-à-vis* ao que se observa nas outras variedades do português brasileiro (PB). Mas não se identificou ainda uma fronteira nítida entre as variedades do PB no que concerne aos valores dos parâmetros sintáticos, sendo esse o primeiro desafio empírico desta investigação, que focaliza a marcação do parâmetro do sujeito nulo. Ao lado desse desafio empírico, colocam-se duas tarefas teóricas. Em primeiro lugar, é preciso definir com segurança como o parâmetro do sujeito nulo se manifesta nas línguas humanas. E, em segundo lugar, como as línguas que se formaram em situações de contato

* Este capítulo é uma versão revista do artigo "Contato entre línguas e variação paramétrica: o sujeito nulo no português afro-brasileiro", publicado na revista *Lingua(gem)* (v.1, n.2, p.63-92), em dezembro de 2004.

linguístico maciço se comportam em relação a esse parâmetro da Gramática Universal (GU). Com base nisso, este capítulo se organiza da seguinte forma: a seção 6.1 traz uma breve apresentação do parâmetro do sujeito nulo no âmbito da Teoria Gerativa; a seção 6.2 discute as manifestações desse parâmetro nas línguas crioulas; já a seção 6.3 faz um apanhado das análises sobre o parâmetro do sujeito nulo no PB; na seção 6.4, são apresentados os resultados de uma análise variacionista da realização do sujeito pronominal em três das comunidades rurais afro-brasileiras isoladas analisadas no âmbito do Projeto *Vertentes*. A seção 6.5 contém as conclusões desta análise.

6.1 O parâmetro do sujeito nulo

O parâmetro do sujeito nulo, o parâmetro *pro-drop*, é o mais discutido e estudado no conjunto da Teoria dos Princípios e Parâmetros, proposta por Chomsky (1981, 1986). Segundo essa teoria, os princípios da Gramática Universal, parte integrante da mente de todos os indivíduos da espécie humana, seriam subespecificados em cada língua particular. Assim, a possibilidade de algumas línguas, como o espanhol e o italiano, de não realizar foneticamente o sujeito pronominal, enquanto em outras línguas, como o inglês e o francês, a sua realização é obrigatória, constituiu o ponto de partida para a proposição do parâmetro do sujeito nulo. Nas línguas *pro-drop*, a possibilidade de recuperar a referência do sujeito através do paradigma flexional de pessoa e número do verbo licenciaria o sujeito nulo, ou seja, uma categoria vazia *pro*. Já nas línguas em que a ausência de flexão verbal não permitisse a recuperação da referência do sujeito, a realização do sujeito pronominal seria gramaticalmente obrigatória. Assim, quanto ao parâmetro do sujeito nulo, as línguas se dividiriam entre aquelas em que o pronome sujeito deve ser sempre realizado foneticamente e aquelas em que esse sujeito pronominal pode ser substituído por uma categoria vazia (configurando-se o *pro-drop*), o que seria licenciado por um paradigma rico de flexão verbal de número e pessoa.[1]

Entretanto, Huang (1989) demonstrou que línguas de tópico desprovidas de um sistema de flexão verbal de pessoa e número, como o chinês, poderiam licenciar o sujeito pronominal nulo. Nesse caso, o licenciamento ocorreria através da correferência com um elemento nominal em uma posição que c-comanda a posição do sujeito pronominal. A questão ainda se torna mais complicada com a

[1] Uma visão mais recente (KATO, 1999b) postula que o próprio morfema de concordância seria um pronominal, inserido como argumento externo do verbo e que depois se afixaria ao I para checar caso nominativo e traços de pessoa e número, não havendo projeção do Spec de I.

notícia de línguas, como o hebraico, que, em alguns tempos verbais e em algumas pessoas, admitem o sujeito nulo, e em outras não (BORER, 1989). Tais achados empíricos vieram a comprometer, em princípio, a relação empiricamente motivada entre a propriedade do sujeito nulo e a riqueza do paradigma flexional de pessoa e número do verbo.

Por outro lado, numa visão mais refinada da teoria dos Princípios e Parâmetros (KATO, 2002a, p. 325), "uma propriedade singular de uma língua não constitui um parâmetro, mas sim a manifestação substantiva de uma propriedade formal abstrata da qual decorrem outras propriedades substantivas na língua", ou seja, cada parâmetro se manifestaria na língua através de um conjunto de propriedades relacionadas. Dessa forma, o parâmetro do sujeito nulo se atualizaria através das seguintes manifestações substantivas relacionadas (RIZZI, 1982):

(i) sujeitos nulos referenciais

(1) a. Fomos à praia. (português)

b. *Went to the beach. (inglês)

(ii) sujeitos nulos expletivos

(2) a. Choveu ontem. (português)

b. *Rained yesterday. (inglês)

(3) a. Parece que vai chover. (português)

b. *Seems that it will rain. (inglês)

(iii) inversão da ordem Sujeito-Verbo

(4) a. Chegaram os hóspedes. (português)

b. *Have arived the guests. (inglês)

(iv) extração do sujeito à distância

(5) a. Quem você disse que veio à festa? (português)

b. *Who did you say that came to the party? (inglês)

Note-se que a correlação de tais manifestações superficiais com uma única propriedade da gramática é muito importante para justificar o poder explicativo e preditivo da teoria gerativista e tem importantes implicações, por exemplo, para a sua teoria da aquisição da língua materna, pois a criança estaria capacitada pelos mecanismos da GU a fazer deduções extremamente eficazes, já que ela não precisa necessariamente ser exposta a todas as propriedades de um parâmetro para fixar o valor da língua-meta. Assim, no escopo da teoria gerativa, a possibilidade de explicar várias propriedades superficiais através de um único princípio da gramática abstrata é mais relevante do que estabelecer uma relação causal entre a riqueza do paradigma flexional e a propriedade do sujeito nulo. Entretanto, estudos da propriedade do

sujeito nulo nas línguas crioulas parecem oferecer contraexemplos a tal generalização.

6.2 O parâmetro do sujeito nulo nas línguas crioulas

Na formalização da Gramática Universal, os parâmetros da variação interlinguística podem ser "concebidos como opções presentes [+ ou 1] ou ausentes [- ou Ø]" (KATO, 2002a, p. 311). De outro modo, pode-se pensar em valores marcados, que dependem de estímulos-gatilhos durante a aquisição da língua materna para serem fixados, ou não marcados, os valores *default* da GU, que correspondem ao estado S_0 da aquisição da língua materna; sendo os parâmetros, portanto, fixados nesses valores na ausência de qualquer estímulo-gatilho. Nesse sentido, pode-se pensar que algumas características estruturais das línguas humanas já estão previstas em S_0, sendo, portanto, inatas ou bioprogramadas (KATO, 2002a, p. 311).

A Teoria do Bioprograma da Linguagem, de Derek Bickerton (1981, 1984, 1988), conjuga-se decisivamente a essa visão. Ao se formarem através da nativização de um modelo precário, de uma língua segunda bastante defectiva (um *pidgin* ou mesmo um jargão, ou pré-*pidgin*), as línguas crioulas tenderiam a exibir um grande número de valores *default* na fixação de seus parâmetros; sendo, entre as línguas naturais, aquelas que mais diretamente refletiriam as propriedades inatas da faculdade da linguagem, ou o estado S_0 da aquisição linguística.

Aplicando esse princípio, Roberts (1997, p. 16-19) define a ausência de sujeitos nulos referenciais como uma opção paramétrica não marcada, comum nas línguas crioulas, na medida em que o licenciamento do sujeito nulo referencial dependeria de um movimento do verbo para Agr-S associado a uma morfologia verbal relevante no que concerne à concordância (ing. *agreement*). No que se refere à possibilidade de essas mesmas línguas crioulas exibirem sujeitos nulos expletivos, Roberts argumenta que o sujeito nulo expletivo não seria uma opção paramétrica, mas uma opção lexical, já que a possibilidade de sujeitos nulos expletivos seria restringida pela presença ou ausência de pronomes expletivos no léxico da língua.

Afirmar que as propriedades da gramática são restringidas pelo inventário lexical, e não o contrário, dificilmente se sustenta, em termos lógicos, ou mesmo frente à teoria do campo. De qualquer forma, o que a reflexão de Roberts destaca é o fato de que as características que as línguas crioulas geralmente exibem em relação às propriedades relativas ao parâmetro do sujeito nulo parecem negar a ideia de um princípio mais abstrato do qual derivam várias propriedades superficiais correlacionadas. Essa parece ser também a conclusão a que se chega com a exposição de Muysken e Veenstra (1994, p. 131-133) sobre o parâmetro do sujeito nulo no

papiamento. Como se pode ver no exemplo (6) abaixo, o papiamento não permite o sujeito nulo referencial, mas permite o sujeito nulo expletivo — cf. (7) —, bem como a inversão e a extração do sujeito — cf. (8) e (9), respectivamente.

(6) a. E ta kome. ('ela está comendo')

b. *Ta kome.

(7) Parse ku Maria ta kanta.

('parece que Maria canta')

(8) Riba e isla aki ta biba um milion hende.

('nesta ilha vivem um milhão de pessoas')

(9) Ken$_i$ b'a bisa k(u)'[e$_i$] a bai festa?

('quem$_i$ você disse que [e$_i$]veio à festa?')

Essa situação do papiamento, que é a mesma do crioulo francês do Haiti e de muitas línguas crioulas, reforça a ideia de que, diferentemente do que a situação do inglês e a do francês sugerem, a impossibilidade de sujeito referencial nulo não implica necessariamente a obrigatoriedade de realização do sujeito expletivo, nem a perda da inversão do sujeito, nem a possibilidade da sua extração. Por outro lado, reforça-se a ideia de que a propriedade de sujeito referencial nulo (*pro-drop*) está fortemente ligada a uma morfologia flexional de pessoa e número do verbo suficientemente forte para permitir a recuperação do sujeito pronominal *pro*. As línguas crioulas de base lexical portuguesa na África, como o cabo-verdiano e o são-tomense, perderam a propriedade do sujeito referencial nulo, exibida pela sua língua lexificadora (o português europeu), mas não perderam a possibilidade de inversão do sujeito e da sua extração, nem gramaticalizaram o uso de qualquer pronome para realizarem o sujeito expletivo, como se pode ver nos exemplos abaixo do cabo-verdiano e do são-tomense.

(10) (Nu bai mar. ↔ *Bai mar. (cabo-verdiano – dialeto de Santiago)

Non ba plé. ↔ *Ba plé. (são-tomense – forro)

'nós fomos à praia'

(11) Tchobe onte. ↔ *Ê tchobe onte. (cabo-verdiano – dialeto de Santiago)

Subê ontê. ↔ *Ê subê. (são-tomense – forro)

'choveu ontem'

(12) Parcen ma ta tchobe. ↔ *Ê parcen ma ta tchobe. (cabo-verdiano – dialeto de Santiago)

Ska ku kwa kuma suba ca ba sobê. ↔ *Ê ska ku kwa kuma suba ca ba sobê. (são-tomense – forro)

'parece que vai chover'

(13) Tchiga ospedes. (cabo-verdiano – dialeto de Santiago)

Chigaza ine convidado. (são-tomense – forro)

'chegaram os hóspedes'

(14) ken ki bu fla ma ben festa? (cabo-verdiano – dialeto de Santiago)

ken gue ku bô fla ku ska bi fesa? (são-tomense – forro)

'quem você disse que veio à festa?'

Fica clara a relação entre a perda da possibilidade de sujeito referencial nulo e um paradigma verbal de flexão de pessoa e número rico, pois, no português dos séculos XV e XVI, havia um morfema flexional para cada uma das seis pessoas do discurso que licenciava o sujeito referencial nulo. A eliminação de toda a flexão verbal de pessoa e número no crioulo cabo-verdiano e no são-tomense implicou a impossibilidade de licenciamento desse sujeito referencial nulo, sem que isso desse início a uma série de mudanças em cadeia que levasse à impossibilidade de inversão e extração do sujeito, e à gramaticalização de um pronome para desempenhar a função de um sujeito expletivo — como seria esperado, segundo as versões mais fortes do parâmetro sujeito nulo. Porém esses crioulos não desenvolveram qualquer uma dessas outras propriedades, pelo menos até hoje, tendo sido decorridos cerca de cinco séculos da sua formação.[2]

Portanto, se o português do Brasil também exibe processos de variação e mudança induzidos pelo contato entre línguas em um grau mais leve do que o observado nas línguas crioulas de base lexical portuguesa da África (cf. capítulo 1, seção 3), pode-se esperar encontrar no Brasil uma diminuição na frequência dos sujeitos referenciais nulos, sem que necessariamente se manifestem outras propriedades correlacionadas, previstas na versão mais forte do parâmetro do sujeito nulo.

6.3 O parâmetro do sujeito nulo no português do Brasil

Analisando os padrões linguísticos da chamada norma culta do português brasileiro, com base em uma amostra de fala espontânea de indivíduos com o curso superior completo, Duarte (1995, p. 141) afirma que os resultados de sua análise variacionista

> [...] revelam que o português brasileiro perdeu a propriedade que caracteriza as línguas de sujeito nulo do grupo *pro-drop* por força do enfraquecimento da flexão, responsável pela identificação da categoria

[2] Assume-se aqui a ideia corrente de que o crioulo cabo-verdiano e o são-tomense se teriam formado entre os séculos XV e XVI.

vazia sujeito em línguas que apresentam uma morfologia verbal suficientemente "rica" para tal processo (ROBERTS, 1993).

Portanto, a análise do preenchimento do sujeito pronominal de referência definida reforça os dois pontos aqui destacados. Por um lado, a perda da propriedade de licenciar o sujeito referencial nulo relaciona-se diretamente com o enfraquecimento do paradigma da flexão de número e pessoa do verbo, pois Duarte (1993) demonstra, em um estudo diacrônico também focalizando a norma culta, que o nível de preenchimento do sujeito pronominal, no seu *corpus* de textos escritos para peças teatrais, salta de 25%, em 1918, para 46%, em 1937; atingindo os níveis de 67% e 74% de sujeitos realizados, em 1975 e 1992, respectivamente. Para a autora, a elevação do nível de realização do sujeito pronominal estaria relacionada com a substituição do pronome pessoal *tu* pelo pronome de tratamento *você* e com a concorrência da expressão *a gente* com o pronome *nós*; processos que ganham corpo nas primeiras décadas do século XX e que levam a uma expressiva redução na flexão verbal, pois, tanto com *você*, quanto com *a gente*, o verbo se mantém na forma da 3ª pessoa do singular, ou seja, sem um morfema específico de pessoa e número. Com isso, em alguns tempos verbais, como no imperfeito do indicativo, a flexão verbal restringe-se a apenas duas pessoas,[3] como se pode ver no Quadro 1. Tal enfraquecimento da flexão verbal estaria, portanto, na base do enfraquecimento da propriedade de licenciar o sujeito referencial nulo no PB.

6 FORMAS	3 FORMAS	2 FORMAS
eu trabalh**o**	eu trabalh**o**	eu trabalhava
tu trabalha**s**	você trabalha	você trabalhava
ele trabalha	ele trabalha	ele trabalhava
nós trabalha**mos**	a gente trabalha	a gente trabalhava
vós trabalha**is**	vocês trabalha**m**	vocês trabalhava**m**
eles trabalha**m**	eles trabalha**m**	eles trabalhava**m**

Quadro 1 - Redução do paradigma da flexão verbal de pessoa e número no português brasileiro

Por outro lado, o PB não perdeu as propriedades de inversão nem de extração do sujeito, nem gramaticalizou um pronome para desempenhar a função de sujeito expletivo, como se pode ver na perfeita aceitabilidade no PB das frases apresentadas em (15), abaixo:

[3] A rigor, o morfema flexional *-m* seria antes um morfema de plural do que propriamente um morfema de pessoa. Portanto, pode-se dizer que, para alguns tempos verbais, o paradigma flexional do verbo de algumas variedades do português brasileiro não possui qualquer morfema de indicação de pessoa do sujeito.

(15) a. Choveu ontem.

b. Parece que vai chover.

c. Chegaram os hóspedes.

d. Quem você disse que veio à festa?

De fato, o PB sequer perdeu a propriedade do sujeito referencial nulo, pois, segundo Duarte (1995), o PB vive uma fase de transição em que convivem características de duas gramáticas. Apesar de ainda admitir perfeitamente frases com o sujeito referencial nulo, a possibilidade de preenchimento do sujeito pronominal em frases como (16), sem mudança de referência ou qualquer ênfase, distingue o PB de línguas tipicamente *pro-drop*, como o italiano, o espanhol ou mesmo o português europeu.

(16) $João_i$ estudou muito, mas ele_i não foi aprovado no exame.

Em sua análise da fala de indivíduos urbanos de alto nível de escolaridade da cidade do Rio de Janeiro, distribuídos pelos dois sexos e divididos em três faixas etárias, Duarte (1995, p. 47) identificou um nível de preenchimento do sujeito pronominal de 71%. Os resultados da variável faixa etária demonstram uma tendência ao incremento do preenchimento do sujeito pronominal; com os falantes mais velhos exibindo os maiores níveis de sujeito nulos referenciais, os falantes mais jovens, os mais baixos. As mulheres, na amostra, lideram a mudança de perda do sujeito nulo (DUARTE, 1995, p. 48-50). No que concerne aos fatores estruturais, os que se mostraram mais refratários ao preenchimento do sujeito foram a 3ª pessoa do singular, em função de sua relação de correferência com um SN anterior, e o sujeito com traço semântico [-animado], pela resistência ao uso do pronome *ele/a* com antecedentes com esse traço semântico. Vale ressaltar que Duarte (1995, p. 40) desprezou as ocorrências de sujeito correferentes em orações coordenadas como exemplificado em (16) acima, por serem contextos em que até línguas não *pro-drop*, como o inglês, admitem o não preenchimento do sujeito, não obstante ter encontrado em sua amostra frases em que o sujeito pronominal é realizado nesse contexto — cf. (17) e (18), abaixo:

(17) Ele trabalhava com a gente aqui na seção de ensino e **ele** está liderando um projeto que ele chamou ‘projeto excelência’.

(18) Eu estranho, mas **eu** fico quieta. **Eu** aceito.

Analisando a língua no plano da sua evolução interna, Duarte (1993, 1995) identifica um processo de mudança em cadeia, em que a mudança na pauta pronominal leva ao empobrecimento da flexão verbal, que, por sua vez, produz uma elevação na frequência de realização do sujeito pronominal. Considerando-se o contexto sociolinguístico da formação do PB, em que se observa uma polarização entre os padrões da fala da elite escolarizada, influenciados pelos modelos do

português europeu, e os padrões da fala da ampla maioria da população, afetados pelo contato entre línguas (LUCCHESI, 1998, 2001a), podem-se agregar novos elementos à compreensão desse processo de perda do licenciamento do sujeito referencial nulo no PB. Em primeiro lugar, se, na norma culta, o móvel da mudança — o enfraquecimento da morfologia verbal — ocorreu em função de substituições na pauta dos pronomes pessoais, na norma popular um enfraquecimento ainda maior da flexão verbal pode ser visto como um reflexo direto do processo de transmissão linguística irregular. Tanto é assim que a redução na flexão verbal, na fala popular, ocorre mesmo quando se mantêm em uso os pronomes *tu* e *nós*, como se pode ver no Quadro 2.

eu trabalho
você ~ tu trabalha
ele trabalha
nós ~ a gente trabalha ~ trabalhamo(s)
vocês trabalham ~ trabalha
eles trabalham ~ trabalha

Quadro 2 - Flexão verbal de pessoa e número no português popular brasileiro

Por outro lado, fica a questão de saber se, como aconteceu em relação aos mecanismos da concordância nominal e verbal, a norma popular apresentará uma tendência diferenciada de mudança, ou exibirá um estágio mais avançado de perda da propriedade de licenciamento do sujeito referencial nulo, aproximando-se mais das línguas crioulas de base lexical portuguesa da África. Essas são algumas das questões que se colocaram para a análise variacionista da realização do sujeito pronominal de referência definida na fala de comunidades rurais afro-brasileiras isoladas, cujos resultados serão apresentados na próxima seção.

6.4 O parâmetro do sujeito nulo no português afro-brasileiro

A análise variacionista da realização do sujeito pronominal que aqui se apresenta tem como base empírica amostras de fala vernácula colhidas em três comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia — Helvécia, Cinzento e Barra/Bananal, no Município de Rio de Contas (cf. capítulo 5, seção 1). A partir dos fundamentos metodológicos da Sociolinguística Variacionista (LABOV, 1972, 1982, 1994), foram eliciadas todas as ocorrências da variável dependente, definida em termos binários pela possibilidade de realização ou não do sujeito pronominal, conforme os exemplos abaixo:

(19) Gente de tempo é assim. Se eu dô o senhô má criaçon, *eles* me dão um côro chateado. Quando [*e*] 'cabá, [*e*] dá o senhô pá surrá ôtra vez.

([*e*] = categoria vazia)

No total de 4.599 ocorrências de sujeito pronominal, em 1.255 ocorrências o sujeito é apagado; o que corresponde a uma frequência de 27 pontos percentuais de realização do sujeito nulo. Em princípio, esse resultado não corresponde exatamente às expectativas iniciais, porque o nível de sujeito pronominal nulo é bastante semelhante ao encontrado por Duarte (1995) na norma urbana culta, como se pode ver na Tabela 1. Entretanto, vale registrar que, por uma divergência metodológica, não foram excluídos da base de dados desta análise as ocorrências de sentenças raízes com verbos epistêmicos tais como "(eu) acho", "(eu) não sei", bem como ocorrências do tipo "sei lá", "sei lá o que" e "não sei o que mais", que Duarte (1995, p. 38-39) retira de sua base de dados, por considerá-las "expressões fixas".[4] Essa diferença metodológica pode ter diminuído a diferença entre o nível de sujeito nulo nas duas variedades do PB.

Tabela 1 - Frequência do sujeito pronominal nulo no português brasileiro urbano culto e no português rural afro-brasileiro

VARIEDADE DO PORTUGUÊS BRASILEIRO	Nº DE OCOR./TOTAL	FREQUÊNCIA
Variedade urbana culta	415/1756	**29%**
Comunidades rurais afro-brasileiras	1255/4599	**27%**

De qualquer forma, o quadro que se insinua é o de que duas variedades opostas do PB, a chamada norma urbana culta e a fala das comunidades rurais afro-brasileiras isoladas, parecem exibir um resultado semelhante em um plano da gramática, a marcação paramétrica, que, segundo a visão aqui adotada, teria sido afetado pelo contato entre línguas; o que constitui, em princípio, uma contraevidência a essa visão, pois o que se espera é que as duas variedades apresentariam comportamentos diferenciados, em função de só as comunidades rurais afro-brasileiras terem sido afetadas diretamente pelo contato maciço e abrupto entre línguas na sua formação.

Por outro lado, pode-se pensar, com base na visão das duas grandes vertentes evolutivas do PB proposta por Lucchesi (1994, 1998, 2001a, 2002b), que houve uma espécie de convergência na evolução gramatical dessas duas variedades linguísticas. Enquanto, na norma culta, a diminuição no nível de sujeito pronominal nulo derivou de uma reorganização estrutural em função das substituições

[4] Por outro lado, observou-se aqui o procedimento de Duarte de excluir as ocorrências em sentenças coordenadas não iniciais com sujeitos correferentes, por serem contextos em que as línguas não *pro-drop*, como o inglês, também admitem sujeitos não realizados.

dos pronomes *tu* e *nós* por *você* e *a gente*, respectivamente; na norma vernácula, essa diminuição estaria ligada diretamente à queda de morfemas flexionais de pessoa e número do verbo, que não seria o resultado de uma reorganização da pauta dos pronomes pessoais sujeito, mas um reflexo direto do processo de transmissão linguística irregular que teria atuado mais intensamente na formação das variedades populares do PB.

6.4.1 Encaixamento linguístico

A análise variacionista do sujeito pronominal nulo de referência definida na fala das comunidades rurais afro-brasileiras isoladas baseou-se no processamento quantitativo dos dados feito pelo pacote de programas VARBRUL (cf. capítulo 5, seção 2). Para o escrutínio do encaixamento linguístico do fenômeno em foco, serão apresentados aqui os resultados de quatro variáveis explanatórias, que o VARBRUL selecionou como estatisticamente relevantes: (i) a pessoa do discurso; (ii) a marca de pessoa e número na forma verbal; (iii) o traço semântico animacidade do sujeito; e (iv) o paralelismo formal.

Os resultados da variável pessoa do discurso, apresentados na Tabela 2, indicam que o sujeito pronominal nulo é mais elevado na terceira pessoa, seguida bem de perto pela primeira pessoa; sendo a segunda pessoa o contexto menos favorável ao sujeito nulo. Esses resultados são semelhantes aos encontrados por Duarte (1995, p. 48), e podem ser explicados pela possibilidade de recuperação do referente do sujeito, seja pelo mecanismo da correferência com um SN antecedente, no caso da terceira pessoa, seja pela presença de um morfema flexional, no caso da primeira pessoa. Assim sendo, o contexto da segunda pessoa, em que, normalmente, não se observa o mecanismo da correferência e que não dispõe de morfemas flexionais específicos no PB, apresenta-se como o contexto mais favorecedor da retenção pronominal.

Tabela 2 - Uso do sujeito pronominal nulo segundo a pessoa do discurso no português afro-brasileiro

PESSOA DO DISCURSO	Nº DE OCOR./TOTAL	FREQUÊNCIA	PESO RELATIVO
1ª pessoa do singular	659/2292	29%	.53
2ª pessoa do singular	27/159	17%	.43
3ª pessoa do singular	392/1167	34%	.55
1ª pessoa do plural	98/671	15%	.34
2ª pessoa do plural	06/35	17%	.43
3ª pessoa do plural	73/275	27%	.49
Total	1255/4599	27%	

Nível de significância: .036.

Essa relação entre o sujeito nulo e a presença de um morfema flexional fica um pouco mais nítida quando se observam os resultados de sua variável específica. Como se pode ver na Tabela 3, o contexto que mais favorece o sujeito pronominal nulo é a presença do morfema *-mos* da primeira pessoa do plural, que é o morfema número-pessoal mais saliente do paradigma flexional, tomando como base aqui o *princípio da saliência fônica* proposto por Naro e Lemle (1976), e já consagrado no campo das análises variacionistas no Brasil. Num nível intermediário, ficam os morfemas da primeira pessoa e o *-m* final, que é a rigor um morfema de número (cf. nota 7). Por fim, o contexto mais refratário ao sujeito nulo é exatamente o da ausência de qualquer morfema flexional. Entretanto, os índices ainda estão muito próximos, e a comparação é feita com base em pessoas do discurso diferentes.

Tabela 3 - Uso do sujeito pronominal nulo segundo a desinência verbal no português afro-brasileiro

DESINÊNCIA VERBAL	Nº DE OCOR./TOTAL	FREQUÊNCIA	PESO RELATIVO
-mos	18/58	31%	.76
-m	20/63	32%	.57
-o / -ei	542/1844	29%	.51
Ø	675/2634	26%	.48
Total	1255/4599	27%	

Nível de significância: .036.

Portanto, a relação entre o sujeito nulo e a presença de um morfema flexional só fica mais nítida quando se faz o cruzamento das duas variáveis analisadas até então. A Tabela 4, que apresenta a frequência de cada pessoa do discurso considerando a presença ou a ausência de seu morfema flexional próprio, confirma a existência da relação entre a ausência de uma marca flexional e a realização do sujeito pronominal. Essa relação é mais significativa com a primeira pessoa do plural, em que a presença do morfema faz o nível de sujeito pronominal nulo aumentar em quase duas vezes e meia, o que se explica, como já dito, pela saliência fônica desse morfema.[5] Esse aumento do sujeito nulo conjugado à presença do

[5] A força do morfema *-mos* pode ser atestada pela sua manutenção nas ocorrências de sujeito pronominal nulo, mesmo entre os falantes do PB que usam normalmente a expressão pronominal *a gente*, para a referência à 1ª pessoa do plural. Tomemos, como exemplo, o seguinte diálogo hipotético:
- *O que vocês fizeram ontem?*
- *A gente saiu.*
- *Foram aonde?*
- *Fomos ao cinema.*
Não seria comum no PB culto a reposta "*Foi ao cinema*". Pode-se dizer que seria quase que agramatical nesse contexto. Quanto ao português afro-brasileiro, está previsto um estudo, no âmbito do Projeto *Vertentes*, que abarque esses e outros aspectos da variação entre *nós* e *a gente*.

morfema flexional reduz-se na terceira pessoa do plural para o patamar de cinquenta por cento;[6] o que pode ser explicado pela possibilidade de recuperar o referente do sujeito, mesmo quando não ocorre a concordância, por meio do mecanismo da correferência com um SN antecedente. Portanto, o que mais surpreende é o resultado da primeira pessoa do singular, em que a presença do morfema não implica em um aumento significativo na frequência de uso do sujeito nulo. Talvez fatores de ordem pragmática e/ou ideológica decorrentes do fato de o indivíduo estar falando de si próprio possam estar intervindo nesse caso; o que fica como hipótese para futuras investigações.

Tabela 4 - Cruzamento da pessoa do discurso com a desinência verbal

PESSOA DO DISCURSO / DESINÊNCIA		Nº DE OCOR./TOTAL	FREQUÊNCIA
1ª pessoa do singular	*-o / -ei*	542/1844	29%
	Ø	117/448	**26%**
1ª pessoa do plural	*-mos*	18/58	31%
	Ø	80/613	**13%**
3ª pessoa do plural	*-m*	20/56	36%
	Ø	**53/219**	**24%**

Os resultados da variável traço semântico animacidade do sujeito também confirmaram os resultados da análise de Duarte (1995). Como se pode ver na Tabela 5, o traço semântico [-animado] favorece de forma considerável a não realização do sujeito de 3ª pessoa. E, como bem observou Duarte (1995, p. 77), o traço semântico [-animado] tem se constituído em um poderoso fator na implementação do uso de categorias vazias no português brasileiro, tanto no caso dos objetos diretos anafóricos (OMENA, 1978; DUARTE, 1989; CYRINO, 1993), quanto no caso do incremento das relativas cortadoras (TARALLO, 1983). Essa mesma tendência se confirma no caso do sujeito pronominal nulo.

Tabela 5 - Uso do sujeito pronominal nulo segundo o traço semântico animacidade no português afro-brasileiro

TRAÇO SEMÂNTICO	Nº DE OCOR./TOTAL	FREQUÊNCIA	PESO RELATIVO
[-animado]	91/165	63%	.73
[+animado]	374/1297	29%	.47
Total	465/1442	32%	

Nível de significância: .036.

[6] As ocorrências da segunda pessoa do plural com o pronome *vocês*, que também se relaciona com esse morfema flexional, foram em número muito reduzido para permitir alguma inferência segura. A segunda e a terceira pessoa do singular não foram consideradas porque, no PB, não existe a possibilidade de virem acompanhadas de qualquer morfema flexional.

Por fim, considerou-se um aspecto do comportamento linguístico dos indivíduos que foi formalizado por Scherre e Naro (1993) através do conceito de *paralelismo formal*. Esse conceito refere-se à tendência de manutenção de um mesmo padrão de marcação no encadeamento da fala. No nível do discurso, o falante tenderia a manter o mesmo padrão de marcação em orações sequenciadas com o mesmo sujeito sintático (SCHERRE; NARO, 1993, p.8-11). Os dados da Tabela 6 demonstram isso, porquanto a presença do sujeito nulo na oração anterior favorece sobremaneira a não realização do sujeito pronominal na oração seguinte (peso relativo de .74). Resta por explicar, no plano psicolinguístico do processamento da fala, por que isso ocorre.

Tabela 6 - Uso do sujeito pronominal nulo segundo o paralelismo discursivo no português afro-brasileiro

POSIÇÃO DA OCORRÊNCIA	Nº DE OCOR./ TOTAL	FREQUÊNCIA	PESO RELATIVO
Antecedido por sujeito pronominal apagado	253/501	50%	.74
Antecedido por SN (expressão referencial)	140/292	48%	.63
Antecedido por sujeito pronominal realizado	297/1171	25%	.50
Primeira referência ou referência isolada	565/2634	21%	.44
Total	1255/4599	27%	

Nível de significância: .036.

Por outro lado, confirma-se também a atuação do mecanismo da correferência no licenciamento do sujeito nulo, pois a presença de um SN antecedente também constitui um contexto favorecedor do sujeito pronominal nulo (peso relativo de .63); do mesmo modo que a primeira referência, onde não há um apoio de um antecedente, constitui o contexto mais refratário ao sujeito nulo (peso relativo de .44). Portanto, confirma-se, através desta análise linguística, os principais fatores que Chomsky (1981) definiu como licenciadores do sujeito nulo: a correferência conjugada à morfologia flexional do verbo.

6.4.2. Encaixamento social

No que se refere ao encaixamento da variação no preenchimento do sujeito pronominal na estrutura social das comunidades rurais afro-brasileiras isoladas, os resultados não proporcionaram um quadro muito nítido. A hipótese inicial seria a de que, no processo de formação dessas comunidades, a violenta erosão da morfologia flexional do verbo teria reduzido em muito a possibilidade de sujeito pronominal nulo. Depois, a influência externa, sobretudo a partir da segunda metade do século XX, teria desencadeado um processo de recomposição da

morfologia flexional do verbo; o que favoreceria, por sua vez, a implementação do sujeito nulo na gramática dessas comunidades. Esse encadeamento estrutural do processo de implementação do sujeito nulo foi confirmado pelos dados da análise linguística, em que se estabeleceu uma relação empiricamente motivada entre o nível de sujeito nulo e a presença de um morfema flexional de número e pessoa na forma verbal. Faltaria demonstrar a tendência de mudança no sentido do incremento do sujeito nulo na gramática das comunidades de fala.

Para testar essa hipótese, foram consideradas na análise quantitativa as seguintes variáveis sociais: sexo, idade dos informantes e seu nível de escolaridade, distinguindo-se aqueles que tiveram qualquer experiência de letramento daqueles totalmente analfabetos. Foi considerada ainda a possibilidade de o indivíduo ter vivido fora da comunidade por pelo menos seis meses. Dessas, o programa das regras variáveis selecionou apenas o sexo e a estada fora da comunidade como estatisticamente relevantes, mas os resultados foram contraditórios. No que se refere ao sexo, os pesos relativos apontam uma ligeira superioridade no nível de sujeito nulo na fala dos homens (.52 *versus* .48, com nível de significância de .036). Isso se ajusta ao raciocínio que tem orientado a análise sociolinguística das comunidades rurais afro-brasileiras isoladas. Por terem mais contato com um mundo exterior, os homens tendem a liderar essas mudanças que refletem uma influência dos padrões de comportamento linguístico exteriores à comunidade; já as mulheres, por estarem mais circunscritas ao universo rural e doméstico, possuiriam uma fala mais conservadora, isto é, uma fala mais distante do padrão urbano culto, ou semiculto.

Entretanto, o resultado da variável estada fora da comunidade vai de encontro a esse raciocínio, pois os seus resultados apontam uma ligeira superioridade na frequência de sujeito nulo na fala daqueles indivíduos que nunca viveram fora da comunidade, praticamente na mesma proporção da diferença entre os sexos (.53 *versus* .48, com o mesmo nível de significância). Ora, segundo a hipótese de que a tendência de mudança seria o incremento do sujeito nulo em função de influências externas às comunidades, o esperado seria o contrário, que aqueles que tivessem vivido fora da comunidade exibissem um nível maior de sujeito nulo na sua fala. Portanto, o quadro não é conclusivo; até porque o principal indicador de um processo de mudança em curso em análises em tempo aparente é a faixa etária (LABOV, 1981, entre outros), e essa variável não foi selecionada pelo Programa das Regras Variáveis. Dessa forma, parece que o quadro que se observa entre a norma culta e vernácula do PB, no que se refere à variação nos parâmetros sintáticos da gramática, é diferente daquele que se observa em relação ao uso da morfologia flexional do nome e do verbo, tanto em termos de suas frequências de uso, quanto em suas tendências de mudança.

6.5 Conclusão

Os resultados da análise variacionista do sujeito pronominal nulo de referência definida na fala das comunidades rurais afro-brasileiras isoladas revelaram uma semelhança entre os condicionamentos linguísticos que regem a variação nessa variedade do PB e na variedade urbana culta. Em ambos os casos, o sujeito nulo é favorecido pela presença de um morfema flexional do verbo, pelo mecanismo sintático da correferência e pelo traço semântico [-animado]. Isso fortalece a visão de uma certa universalidade da ação dos condicionamentos estruturais sobre os processos de variação e mudança no sistema da língua, tanto no plano diacrônico de suas variedades linguísticas, como proposto por Kroch (1994), quanto no plano sincrônico de suas diversas variedades diastráticas, como se observou aqui. Por outro lado, no plano sociolinguístico, constatou-se uma similaridade na frequência de uso do sujeito pronominal nulo no padrão urbano culto e no padrão vernáculo das comunidades afro-brasileiras.

Com base na visão de uma formação polarizada do PB desenvolvida por Lucchesi (1994, 2001a, 2002a, 2002b, 2006a) pode-se postular um desenvolvimento paralelo, e possivelmente convergente, em que na norma culta estaria em curso uma diminuição na frequência de uso do sujeito nulo desencadeada por um processo de reorganização na pauta dos pronomes sujeito, ao passo que, nas diversas variedades populares, essa propriedade já teria sido restringida em graus diferenciados, consoante o nível de erosão de sua morfologia flexional, determinado pela intensidade diferenciada dos processos de transmissão linguística irregular em sua formação. Desse modo, os padrões de uso culto do PB estariam se aproximando do nível de variação paramétrica já assente em suas variedades populares; o que poderia explicar as tendências convergentes de mudança.

Diferentemente do que ocorre com a morfologia flexional, em que as variedades populares do PB apresentam uma tendência de incremento do uso da morfologia flexional (ou seja, uma mudança que elimina os efeitos do contato entre línguas nesses dialetos), no que concerne à variação paramétrica, não se observa um quadro de incremento do sujeito nulo. Assumindo a hipótese de que a forte erosão da morfologia flexional ocorrida na formação desses dialetos teria restringido bastante a possibilidade do sujeito pronominal nulo de referência definida, seria de esperar que, com a recomposição da morfologia flexional, ocorresse simultaneamente um incremento do sujeito nulo. Um fator que pode estar comprometendo essa relação é a reação subjetiva dos falantes ao processo de variação e mudança na realização do sujeito pronominal. Essa variável não constitui um traço estigmatizado, no plano da avaliação social da variação linguística, como o é a falta da concordância nominal e verbal. Esse sensível fator subjetivo que estaria impulsionando a implementação do uso das regras de concordância nas variedades

populares do PB — como previsto nas formulações de Weinreich, Labov e Herzog (1968) e Labov (1982) — não estaria atuando no sentido de implementar o sujeito pronominal nulo. Ao lado disso, a norma culta não exibe um nível robusto de sujeitos nulos que possa servir de modelo para os falantes das classes mais baixas; ao contrário do que ocorre em relação à concordância nominal e verbal, em que os usos dos morfemas de plural da segunda e da terceira pessoas verbais e dos nomes e adnominais guardam uma grande vitalidade nos usos cultos monitorados, constituindo um modelo de comportamento linguístico para toda a sociedade. Além de tudo, as variedades populares têm acompanhado, ou mesmo superado, a norma culta na substituição do pronome *nós* pela expressão pronominal *a gente* (cf. capítulo 19), reduzindo sobremaneira o emprego do morfema *-mos* da primeira pessoa do plural — exatamente o morfema que mais favorece o sujeito nulo, conforme demonstrado pelos resultados quantitativos desta análise.

Tais fatores podem estar se conjugando para sustar uma tendência de incremento do sujeito nulo nas variedades populares do PB, fazendo com que, a partir de motivações diversas, em todas as variedades do PB, se configure um quadro de incremento da realização do sujeito pronominal que pode vir a levar à perda da propriedade do sujeito nulo de referência definida no PB. De qualquer forma, a complexidade do quadro que se delineia quando se conjuga a variação paramétrica ao caráter polarizado da realidade linguística brasileira e aos processos de contato linguístico que concorreram para a sua formação ressalta a necessidade de novas investigações empíricas para esclarecer as correlações que existem entre tais fatores sócio-históricos e os processos estruturais que estão dando forma à gramática das diversas variedades da língua portuguesa no Brasil.

7 As sentenças relativas

Ilza Ribeiro

> Se ele enfezá, ele não sai, não. É... ocê tem que ir com cuidado, que, no lugá onde ocê tá, ir no lugá adonde ele tá,[...]. Cê vem com sentido onde ele tá e... pisa no fucinho, no lugá donde cê tava. Sai do lugá dond'ele tava e vai pro lugá dond'a pessoa tava. (RC-08) [1]

Diversos estudos, quer de abordagem diacrônica (TARALLO, 1983, 1993a, 1993b; COHEN, 1986; JESUS, 2002), quer de abordagem sincrônica (LEMLE, 1978; KATO, 1993; KATO et al., 1996) têm discutido as características das sentenças relativas no português brasileiro (PB). Grosso modo, as discussões se centram em três propriedades fundamentais das estratégias relativas não padrão: (a) a ausência de preposição antecedendo o pronome relativo (PR) nas relativas de funções argumentais preposicionadas e de adjunto, denominadas relativas *cortadoras*; (b) a possibilidade de presença de pronome lembrete nas posições relativizadas, nas chamadas relativas lembrete ou *resuntivas*; (c) o estatuto gramatical do relativizador *que*, um complementador ou um PR.

As pesquisas sincrônicas têm estabelecido comparações entre o PB culto e o português europeu (PE) culto, bem como entre diversos dialetos regionais do PB, entre fala e escrita, entre falantes escolarizados e não escolarizados e outros fatores condicionadores de variação social (LEMLE, 1978; KATO, 1981; KATO et al., 1996; CORRÊA, 1998, 1999). As diacrônicas apontam para mudanças no PB na virada do século XVIII para XIX, momento em que as propriedades (a) e (b) se mostram quantitativamente mais produtivas (TARALLO, 1983, 1993a, 1993b; JESUS, 2002).

O objetivo deste capítulo é discutir a sintaxe das sentenças relativas na fala de comunidades rurais afro-brasileiras isoladas, comparando seus resultados com os de pesquisas já realizadas sobre aquisição de relativa em primeira língua (L1),

[1] Exemplo retirado do *corpus* em análise. A sigla indica a comunidade: RC representa Rio de Contas. As outras são: CZ, Cinzento ; HV, Helvécia e SP, Sapé. As comunidades são descritas no capítulo 2 deste livro, e a metodologia empregada na constituição das amostras de fala vernácula é apresentada no capítulo 5.

em fala do PE e de crioulos de base lexical portuguesa. A análise apresentada neste capítulo contempla as estratégias de relativização atestadas nos *corpora* constituídos a partir dos inquéritos de 8 informantes das seguintes localidades do interior da Bahia: 1) Cinzento, informante 06, 48 anos, semianalfabeto (CZ-06); 2) Cinzento, informante 08, 50 anos, analfabeto (CZ-08); 3) Helvécia, informante 13, 85 anos, analfabeto (HV-13); 4) Helvécia, informante 20, 70 anos, analfabeto (HV-20); 5) Rio de Contas, informante 08, 55 anos, analfabeto (RC-08); 6) Rio de Contas, informante 26, 68 anos, semianalfabeto (RC-26); 7) Sapé, informante 09, 76 anos, analfabeto (SP-09); 8) Sapé, informante 12, 66 anos, analfabeto (SP-12).

Do ponto de vista teórico, assumo que só há uma forma de construir a gramática de uma língua, através das restrições impostas pelos princípios universais inatos, seja em aquisição de L1, de L2 ou de crioulização; do ponto de vista empírico, o estudo das sentenças relativas nessas comunidades não fornece evidências de um processo anterior de pidginização/crioulização do PB (cf. discussões em GUY, 1989; TARALLO, 1993, 1993a, 1993b; LUCCHESI, 2000a), nem da hipótese da deriva (NARO; SCHERRE, 1993); mas argumenta a favor da transmissão linguística irregular (LUCCHESI, 2000a, 2003b, 2008c).[2]

O texto procura responder a questões tais como:

(i) Como se dá a relativização neste *corpus*, em relação às estratégias, ao encaixamento e às funções sintáticas dos marcadores de relativa? Como analisar esses dados em relação às restrições universais para a formação de sentenças relativas, segundo estudo de Keenan e Comrie (1977, 1979)?

(ii) Qual é o comportamento das relativas cortadoras e resuntivas neste *corpus*? Esses dois tipos de relativas podem ser considerados indícios de transmissão linguística irregular? (Cf. discussão em TARALLO, 1993a, 1003b; NARO; SCHERRE, 1993; LUCCHESI, 2000a).

(iii) As estratégias em uso são semelhantes ou diferentes das observadas na aquisição de outras L1?

(iv) Quais tipos de estratégias são atestados em crioulos de base portuguesa? Os informantes do *corpus* em estudo realizam estratégias semelhantes?

(v) Tem-se observado que em processos de transmissão linguística irregular há sempre perda / reanálise de morfologia. Os tipos de marcadores de relativa usados pelos informantes indicam perda / reanálise morfológica?

O capítulo está organizando em torno dessas questões, portanto. Na seção 7.1, são definidos e apresentados os tipos de relativa dos *corpora* em estudo (relativas restritivas, explicativas e relativas livres ou semilivres); na seção 7.2, discutem-se as

[2] O conceito de *transmissão linguística irregular* também é objeto do capítulo 3 deste livro.

questões sobre restrições universais no que diz respeito ao encaixamento da sentença e às funções sintáticas do PR; também são analisadas as estratégias de relativização (lacuna, cortadora, lembrete e *pied piping*); uma discussão sobre PR e complementador é o tema da seção 7.3, com o objetivo de esclarecer a hipótese central do trabalho, a de que, nessas comunidades, as relativas com antecedente são sempre introduzidas pelo complementador *que* e os PRs são especializados para as relativas livres. Os dados que sustentam essa hipótese são apresentados nas subseções de 7.3. Fechando o capítulo, na seção 7.4, argumenta-se contra a hipótese da deriva, contra evidências de crioulização prévia e a favor da transmissão linguística irregular.

7.1 Tipologia das sentenças relativas

Uma sentença relativa é formada por um núcleo nominal modificado por uma sentença. Por exemplo, o DP sujeito da sentença em (1)

(1) *[[O livro]$_{DP1}$ [que eu comprei ___]$_{SRel}$]$_{DP2}$*

consiste de um determinante (*o*), de um núcleo nominal (*livro*) e de uma sentença que modifica o núcleo nominal (***que** eu comprei* ___); a sentença modificadora é introduzida por um relativizador (*que*). Observa-se que, na sentença relativa (SRel), o objeto do verbo *comprar* não está realizado, sua posição estando marcada por ___, significando uma lacuna. O núcleo nominal *livro* é entendido como objeto de *comprar*; todo o DP que contém o núcleo nominal *livro* é sujeito do predicado *discutir*. Assim, o núcleo nominal parece realizar, ao mesmo tempo, duas funções gramaticais. Esse tipo de relativa com lacuna tem sido denominada *relativa padrão, relativa com lacuna, relativa convencional*. Adoto aqui a denominação mais usual, *relativa com lacuna* (cf. seção 7.2).

7.1.1 Sentenças relativas restritivas e não restritivas

O contraste entre relativas restritivas e relativas não restritivas (ou apositivas) pode ser ilustrado com os exemplos em (2):

(2) a. A diretora da escola está procurando [*o menino* que fugiu da sala de aula ontem]

b. A diretora da escola está procurando [*Pedrinho*, que fugiu da sala de aula ontem]

Na relativa restritiva em (2a), a SRel restringe a classe de *menino* que a diretora está procurando, referindo a somente aquele *que fugiu da sala de aula ontem*. Assim, a referência específica do DP *o menino* é determinada pelo modificador clausal. Em (2b), a relativa não restritiva serve como um comentário à parte, separado

do núcleo nominal *Pedrinho* com entonação adequada, marcada na escrita por vírgula. O referente do núcleo nominal já é conhecido ou pode ser identificado independentemente da SRel. Desse modo, a relativa não restritiva simplesmente apresenta informação adicional sobre o núcleo nominal.

Os dois tipos de relativas são atestados no *corpus* em estudo, perfazendo o total de 82% dos dados coletados. Contudo, são as relativas restritivas que predominam nos dados, com 65% de ocorrências (exemplos em 3), sendo 17% de não restritivas (exemplos em 4). Os 18% restantes da base de dados correspondem às relativas livres e semilivres, que serão apresentadas na seção 7.1.2., abaixo.

(3) a. no lugá *qu'*eu morava, na fazenda *qu'*eu morava... (SP-09)

b. tipo de...de...de raízes *que* usava pra determinado caso (RC-08)

c. Todos negóço *qu'*eu quero fazé e coisa ... (SP-12)

(4) a. lá no faiado, *onde* Dona Maria mais Irto morava ... (HV-13)

b. Meu pai morreu, inda tem minha mãe, *que* mora comigo, (CZ-06)

c. Trabaio na minha rocinha, *qu'*eu tenho, e trabaio ganhano o dia. (SP-12)

No *corpus* em estudo, as relativas restritivas e não restritivas são quase categoricamente realizadas com o relativizador *que* (93%); foram atestadas só 16 ocorrências com *onde* (3%) e uma ocorrência de *quando* (ex. em 5c).[3]

Há algumas ocorrências de relativas não restritivas 'extrapostas'; portanto, separadas do seu núcleo nominal:

(5) a. Benvino morreu, *qu'*é que me 'companhô... *qu'*é mais velho de qu'eu. (HV-13)

b. o mais véi foi... tava com doze ano, *que* era Tecílo, tava com doze ano...quano meu pai morreu... (CZ-08)

c. naquele tempo era bifase, *quando* eles botô (CZ-06)

Os exemplos em (5) são construções inacusativas.[4] Isso significa dizer que os constituintes pré-verbais (*Benvino*, *o mais véi*, *naquele tempo*) foram alçados de uma posição interna aos predicados, deixando a SRel nesta posição básica interna.

7.1.2 Relativas livres e semilivres

As *relativas livres* se caracterizam pela ausência de um antecedente explícito para a referência do constituinte relativo que a contém, por isso também denominadas *relativas sem antecedente*. Assim, as relativas livres são necessariamente

[3] Há 18 ocorrências de *o que* variável (4%) que foram analisadas como relativas semilivres. Cf. item a seguir.

[4] Sobre inacusatividade na fala destas comunidades, cf. Carvalho, 2006.

introduzidas por um PR, nunca por um complementador nulo, nem por um relativizador tipo *that* do inglês ou *que* do francês.

(6) a. *Onde* João estava foi na minha casa

b. *Quem* comprou o livro foi João

c. *O que* eu comprei foi este livro

d. *Quando* eu fui a Paris era primavera

e. *Como* eu fiz isto foi desse modo

As *relativas semilivres* ocorrem com as formas *o que* e variações morfológicas, apresentando um nome nulo como antecedente do constituinte relativo; o determinante *o* concorda em gênero e número com este nome nulo. Exemplos dos dois tipos são dados a seguir:

(7) a. *O que* eu li foi estas revistas

b. *As* que eu li foram estas revistas

O PR *o que* é invariável em (7a); contudo é variável em (7b), concordando com o núcleo nominal nulo (*as revistas*).

As relativas livres (exs. em 8) e semilivres (exs. em 9) totalizam 18% dos dados em estudo:

(8) a. *Quano* é nesse tempo d'agora...d'agora pra janêro, aí chega... *o que* tem de chegá... chega. (RC-26)

b. graças a Deus... nóis reza *o que* minha mãe me ensinô, (CZ- 08)

c. *Quem* não faz justiça tem processo! (HV-20)

d. É, *onde* tem a casa, *onde* tem sua morada, *onde* tá parente... parente tá em qualqué lugá, mas *aonde* cê é mais acostumado é *ondé* que cê tem mais parente, né? (CZ-06)

As formas pronominais que se realizam nas relativas livres (103 ocorrências) se distribuem como segue: *quem* (44%), *onde* (31%), *o que* (17%) *quando* (8%).

Foram atestados 18 dados de relativas semilivres, como ilustrado em (9):

(9) a. *A que* celebô no dia de Santa Luzia foi aquela desse dia que cê tava aqui (CZ-06)

b. *Os que* tava aqui vêi, ficarum desempregado lá (CZ-06)

c. *As que* é feia, eu tranco o rádi... (CZ-08)

d. *A que* tava era de Bernardino, que era o mais velho. (HV-20)

Os demonstrativos das relativas semilivres alternam entre *o* e *esse*, *aquele*:

(10) a. *aqueles que* vai de... de premêra recebe, *os que* ficá pá ir de derradêra num recebe mais (CZ-06)

b. Mas, com a fé de *Deus*, *esse que* tem de sarvá, sarva. (HV-13)

7.1.3 Relativas existenciais

Do total de dados coletados, 19,3% são realizações de *relativas existenciais*:

(11) a. *tem* um cara ni Planato *que* é deputado lá. (CZ-06)

b. Às vez, tem um doente que fazia pomessa, que meorava... (HV-13)

c. aí tinha uma fia que morava lá em Cajaíba... (SP-12)

d. Não, aí embáxo *tem* um home *que* faz... *que* vende aí tamém. (SP-09)

Há também 13 casos de relativas com *ter* de posse; observa-se que em um mesmo enunciado pode haver variação entre *existência* e *posse*, como os dados a seguir ilustram:

(12) a. e *tenho* uma irmã *que* mora no Maitá, no municípi de... de Anagé e *tem* ôto *que* mora no Pendanga, (CZ-06)

b. Eu *tenho* muitos parente aqui *que* mudô daqui den'dessa comunidade, (CZ-06)

c. *tem* um *que* é pai de família... (CZ-06)

d. eu *tenho* um *que* morava em Conquista, (CZ-06)

7.1.4 Conclusões parciais

Outras formas do PR atestadas em outras variedades do PB estão ausentes neste *corpus*. Por exemplo, não há ocorrência de *quanto*, embora o contexto para sua realização seja atestado:

(13) a. *tudo que* Deus dá a gente tem que recebê aquela... aquela graça que Deus dá, (SP-12)

b. e ela queria tomá *tudo que* eu teve, mas aí a justiça num entregô. (HV-20)

Em outras variantes dialetais do PB, o pronome *quanto* pode ocorrer, em realizações como as em (14):

(14) a. *Tudo quanto* Deus nos dá ...

b. *Tudo quanto* eu tive ...

De modo semelhante, em contextos em que se pode realizar o pronome *como*, só se observa neste dialeto o uso de *que*:

(15) a. *do jeitchim que* nós tinha feito... do tempo de meu pai, do meu avô... (CZ-06)

b. *Do jetcho que* eu dêxei lá, eu num vi não... (CZ-08)

Também não foram atestadas ocorrências de *o qual*, de uso muito frequente no PB, com diferentes funções:

(16) a. o professor, o qual não gosta de mim, sempre me dá nota baixa.

b. o professor, para o qual você enviou o trabalho, sempre me dá nota baixa.

7.2 Sentenças relativas e restrições universais

Keenan e Comrie (1977, 1979), com base nos resultados de análise de um número amplo de línguas de diferentes famílias linguísticas, definem restrições universais para a formação de SRel e identificam três estratégias mais comuns de relativização, de maior uso nas línguas estudadas, grosso modo como denominadas neste texto: (a) a estratégia da *lacuna*; (b) a do *pronome resuntivo*; (c) e a estratégia do *pronome relativo*.

(17) a. [A sala$_i$ [*que* a janela —-$_i$ está quebrada$_{SRel}$] fica no corredor à esquerda

b. [A sala$_i$ [*que* a janela *dela*$_i$ está quebrada$_{SRel}$] fica no corredor à esquerda

c. [A sala$_i$ [*cuja*$_i$ janela — está quebrada$_{SRel}$] fica no corredor à esquerda

A designação da estratégia em uso no exemplo (17a) como *estratégia de lacuna* não é muito feliz, pois em (17c) também há uma lacuna na SRel. Contudo, a escolha de SRel de genitivo para ilustrar a classificação permite identificar facilmente as três estratégias, sobretudo no que diz respeito às que deixam sempre uma lacuna dentro da SRel, diferindo em que (17a) ilustra o uso de um complementador, enquanto (17b), o de um PR.

Os autores observam que não há variação aleatória quanto ao tipo de posição sintática do NP relativizado, ou seja, em relação à função do elemento relativizado na SRel. A hierarquia proposta segue a seguinte cadeia (KEENAN; COMRIE, 1977, p.66):

(18) Hierarquia da Acessibilidade (Accessibility Hierarchy; doravante AH)

SU > OD > OI > OBL > GEN > Objeto de Comparação[5]

A AH de funções relativizadas em (18) reflete o fato de que algumas posições são mais acessíveis à relativização do que outras, ou seja, as posições à esquerda da AH são mais fáceis de relativizar do que as posições à direita. A generalização derivada da AH é que, se uma língua pode construir SRel de uma dada posição da AH, então pode também construir relativas de todas as outras posições à sua esquerda. Assim, se uma língua pode relativizar objetos, também pode relativizar sujeito; qualquer língua que possa relativizar genitivos também pode relativizar oblíquos e as demais funções à esquerda.

Das predições feitas por Keenan e Comrie (1977, 1979), as seguintes são relevantes para este trabalho:

(19) a. Nenhuma língua pode construir relativas somente de objetos ou somente de locativos, mas é possível uma língua construir relativas somente de

---

[5] Keenan e Comrie (1977, p. 74) apresentam o seguinte exemplo de relativa de objeto de comparação: *The man who Mary is taller than*. Para Kato (198, p. 2), a relativa de objeto de comparação é possível em PB com a estratégia do pronome resuntivo: *?O homem que meu pai tem mais força que* ***ele***.

sujeito (este é o *Subject Relative Universal* (KEENAN; COMRIE, 1979, p. 652). Assim, todas as línguas podem relativizar sujeito.[6]

b. Uma língua deve ter uma estratégia primária (*primary strategy*) de formar SRel, quer dizer, qualquer uma das estratégias referidas acima (lacuna, pronome resuntivo, pronome relativo) pode ser a estratégia primária.

c. A estratégia primária pode deixar de ser realizada em qualquer ponto da AH, ou seja, estratégias que se aplicam a um ponto da AH podem não se aplicar a qualquer posição à direita. Consequentemente, se a estratégia primária, em uma dada língua, pode ser aplicada a uma posição baixa, então pode também ser aplicada a todas as posições mais altas da AH.

7.2.1 Função sintática do pronome relativo na amostra de fala analisada

Os dados da comunidade apresentam a seguinte hierarquia, no que diz respeito à função sintática relativizada:

(20) SU(51,2%) > OBL(23,8%) > OD(15%) > OI(8,7%) > GEN(0,5%)

Em 51,2% das construções, a função sintática relativizada é a de sujeito; em 15%, a de objeto. As únicas três construções codificadas como de genitivo são:

(21) a. Mas antigamente... era uma casa *que* a gente oiava d'um lado tava veno o ôtro... (SP-12)

b. Ieu só vejo os ôto falá. Tem muita gente *qu'eu* num conhece ninguém... (HV-13)

c. Ieu mesmo tenho uns... um minha... uma sobrinha *que* [o] marido *dela*... que tem roça pra lá, no faiado... dá pra lá. (HV-13)

considerando as leituras:[7]

(22) a. gente olhava de um lado *da casa*

b. não conheço nenhuma *destas pessoas*

c. o marido *da sobrinha*

No cômputo geral, portanto, há mais SRels com foco no sujeito, estratégia que parece ser a mais recorrente nas línguas humanas, sobretudo na fase de aquisição da linguagem (ROMAINE, 1988; GUASTI; CARDINALETTI, 2003). A hierarquia

---

[6] Se a língua permite relativização, pois uma língua pode não ter desenvolvido qualquer estratégia de relativização.

[7] Evidentemente, outras leituras são possíveis. Para (21a), por exemplo, Dante Lucchesi (c.p.) sugere a seguinte leitura: (i) era uma casa *que* a gente oiava **ela** d'um lado tava veno o ôtro...

em (20), contudo, difere da AH em (18), em que há mais funções relativizadas de oblíquo do que de objeto direto. Uma explicação plausível para esta diferença pode ser formulada em termos de um viés do *corpus*, pois os documentadores fazem muitas perguntas sobre ONDE e QUANDO (onde planta, mora, nasceu.... quando planta, colhe, chove...).

7.2.2 Encaixamento da relativa

A estratégia mais frequente é a de SRel encaixada em constituinte nominal que exerce a função de objeto, de verbo (37,2%) ou de preposição (9,4% para OI e 12,6%, para adjunto adnominal e complemento nominal), o que permite o encaixamento na posição mais à direita do constituinte, facilitando o seu processamento (SLOBIN, 1986). O encaixamento na função de sujeito é maior (23,8%) do que na dos adjuntos oblíquos (17%).

(23) OD(37,2%) > SU(23,8%) > OBL(17%) > ADN/CN(12,6%) > OI(9,4%)

Parece haver uma tendência nas línguas humanas a realizar mais o encaixamento na posição de objeto do que na de sujeito. Os resultados de Romaine (1988) sobre aquisição de relativas por falantes de inglês em fase escolar (*Edinburgh*), como L1, mostram que as crianças realizam relativas com 71% de encaixamento no objeto, contra 29% no sujeito. Ribeiro e Figueiredo (2006) também apontam, como resultados do estudo das realizações de relativas em atas escritas por africanos livres, no Brasil, no século XIX, o maior percentual de encaixamento na posição de objeto (87,5%) do que na de sujeito (12,5%).

7.2.3 Estratégias de realização das relativas

Nesta subseção faz-se uma apresentação geral das estratégias mais comuns de relativização nas línguas humanas; tem-se como objetivo principal estabelecer comparações entre as possibilidades presentes em diversas línguas e os usos atestados nas comunidades em estudo.

7.2.3.1 Relativa *pied piping* vs relativa cortadora

A relativa *pied piping* é um tipo de relativa com lacuna, específica das posições preposicionadas. Em inglês, a relativa *pied piping* se opõe à relativa com preposição encalhada (*stranding*); em português, a oposição se dá em relação à cortadora (exemplos em (24a, b) e (24c, d), respectivamente:

(24) a. Jack is the boy *with whom* Jenny fell in love (*pied piping*)

b. Jack is the boy *that* Jenny fell in love *with* (preposição encalhada)

c. A moça *com quem* conversei ontem (*pied piping*)

d. A moça *que* eu conversei ontem (cortadora)

Não foi observado, na amostra de fala analisada, o uso de relativa com antecedente na estratégia *pied piping*, estratégia que tem se tornado rara no PB em geral. Todos os casos de funções preposicionadas se realizam na estratégia cortadora, ou seja, sem a preposição, como nos exemplos:

(25) a. mandô aí, a muié de lá de Salva... de Jequié, essa Rosa, *que* eu tô falano, veio aqui olhô... (CZ-06)

b. Num adianta insisti... num caminho *que* não tem saída, (CZ-06)

c. No dia mesmo da missa *que* teve sete pade aqui mesmo, (CZ-06)

d. no lugá *que* dá ôro grosso, a pessoa vai garimpano...(RC-26)

A relativa cortadora é a estratégia presente nos dialetos do PB em geral; a estratégia *pied piping* é um recurso adquirido via escolarização, evidenciando uma situação de bilinguismo, no sentido de Kroch (2005).

Contudo, a relativa cortadora não é uma particularidade do PB, em oposição a outras línguas românicas. Também no francês e no italiano, a estratégia *pied piping* está mais restrita à fala formal e à escrita, caracterizando-se como um artefato prescritivo, que emerge como resultado de ensino explícito, durante a escolarização (GUASTI; CARDINALETTI, 2003). No PE culto contemporâneo, os estudos de Alexandre (2000) e de Arim, Ramilo e Freitas (2004) apontam para a tendência à variação no uso das duas estratégias, como em:

(26) a. Isso é um atributo *que* as pessoas, prontos, gostam. (Dinheiro Vivo, RTP2) (ARIM; RAMILO; FREITAS, exemplo (9))

b. A rapariga *de que* eu gosto faz-me rir. (ARIM; RAMILO; FREITAS, exemplo (17))

7.2.3.2 Relativa com pronome lembrete

São poucos os casos atestados de relativa com pronome *resuntivo*:

(27) a. Mas teve um prefeito... um prefeito... aí *qu'*eu gostei *dele*, foi dotô Pedro (RC-26)

b. Tem um camim, *que* pega ôndibu *lá*, vai embora po Texêra, (HV-13)

c. Essa lagoa... lugá *que* Alícia mais Ilton mora *lá*, (HV-13)

d. cê topa uma pessoa lá *que* cê trabaia dereitin *com ele* (CZ-06)

e. botava naquela estrada *que* cês passa... cês passa *nela*... (CZ-08)

As poucas ocorrências de resuntivo dizem respeito, basicamente, às retomadas pronominais com advérbios locativos (27b, c, e).

Chama atenção a falta de *resuntivo* por ser esta uma estratégia bastante atestada nos dados de diferentes línguas, como ilustra a Tabela 1.

Tabela 1 - Perspectiva translinguística de relativas com pronome *resuntivo*

ESPANHOL		IÍDICHE	IRLANDÊS	GALÊS	HEBRAICO	PALESTINO
SU	∅ / Pron.	∅ / Pron.	∅	∅	∅	∅
OD	∅ / Pron.	∅ / Pron.	∅ / Pron.	∅	∅ / Pron.	Pron.
OB	∅ / Pron.	Pron.	Pron.	Pron.	Pron.	Pron.
GEN	∅ / Pron.	Pron.	Pron.	Pron.	Pron.	Pron.

Fonte: adaptada de SUÑER, 1998, p.49.

Notas: ∅ = lacuna na posição sintática vinculada ao pronome relativo; Pron. = realização do pronome resuntivo.

O espanhol e o iídiche realizam a estratégia com pronome resuntivo para todas as funções; o irlandês, o hebraico e o palestino só recusam a estratégia resuntiva para a função de sujeito; o galês é a língua mais restritiva, só aceitando o pronome resuntivo para as funções de oblíquo e genitivo.

Vários estudos têm apontado a quase ausência da estratégia resuntiva no que diz respeito ao PB, quer na fala de adultos, quer em dados de aquisição (TARALLO, 1993a, 1993b; CORRÊA, 1998; PERRONI, 2001; BAGNO, 2001; BURGOS, 2003, dentre outros). Também no PE contemporâneo, esta não é a estratégia mais em uso, como mostram os resultados quantitativos apresentados na Tabela 2.

Tabela 2 - Estratégias de relativização no PE

TIPO DE RELATIVA	CASOS ATESTADOS	%
Relativas canônicas	189	71%
Relativas cortadoras	74	28%
Relativas resuntivas	2	1%
Total	265	100%

Fonte: Arim, Ramilo e Freitas, 2004.

As relativas canônicas, ou com lacuna, perfazem 71% dos dados do PE. Há só dois casos de relativa resuntiva (1%), indicando que esta não é a estratégia preferida; os 28% de relativa cortadora apontam para um uso regular deste tipo de relativa.

7.2.4 Conclusões parciais

A seção 7.2. procurou realçar os principais fatos que nortearão as análises das seções 7.3 e 7.4, a saber: (i) o encaixamento da relativa e a função sintática relativizada seguem a AH proposta por Keenan e Comrie (1977), como era esperado, desde que a AH deve se refletir em qualquer dialeto; (ii) ausência total de relativa

pied piping, sendo a relativização de funções preposicionadas realizada pela estratégia cortadora; (iii) pouco uso da relativa lembrete, indicando não ser esta a estratégia preferida, como também acontece no PB em geral e no PE.

Embora as estratégias cortadora e lembrete tenham, durante muito tempo, sido consideradas um divisor de gramáticas entre o PB e o PE, recentemente estudos sobre os processos de relativização no PE têm apontado para a existência, também lá, das estratégias cortadora (cf. exemplos em (28a, a')) e lembrete (cf. exemplos (28b, b')), quer nos meios de comunicação social (ALEXANDRE, 2000; ARIM; RAMILO; FREITAS, 2004),[8] quer no dialeto rural (VAREJÃO, 2006).[9] Seguem-se alguns exemplos ilustrativos (cf. também exemplos em (26)):

(28) a. Não fixei o nome da pessoa **que** tu falaste. [*Conversas Secretas*, SIC] (Arim, Ramilo e Freitas, 2004, exemplo (12)).

a'. Há comeres **que a gente também enjoa** (Alcochete 4) (ex. (144) de VAREJÃO, 2006).

b. Sei de *um caminho*$_i$ [[$_{OBL}$ *que*] o pai passou *por lá*$_i$ da outra vez]. (ALEXANDRE, 2000, Anexo I, 2 (26)).

b'. Há cardumes de tainhas, há cardumes de robalos — aqui mesmo na costinha — que **uma pessoa vê-os.** E dizem assim: "Ai rapaz! Que cardume de robalos!" (Vila Praia de Âncora 24) (ex. (145) de VAREJÃO, 2006).

Segundo Varejão (2006), do total de 103 ocorrências do *corpus* do PE rural, 75 (73%) apresentavam o corte da preposição e 28 (27%) apresentavam cópia. Além disso, a estratégia *pied piping* é atestada:

(29) a. Parece que isto que é talvez o dote **com que** nasce as pessoas (Cabeço de Vide 65) (exemplo (164) de VAREJÃO, 2006)

b. Depois eram atiradas para aquele lugar **em que** ficava com um coiso assim (Cabeço de Vide 34) (exemplo (165) de VAREJÃO, 2006)

c. Isto é os raios **com que** eles agarram (Câmara de Lobos 26) (exemplo (166) de VAREJÃO, 2006)

d. Houve uma altura **em que** nevou muito (Enxara do Bispo 15) (exemplo (167) de VAREJÃO, 2006)

[8] Arim, Ramilo e Freitas (2004) trabalham com dados de textos veiculados nos meios de comunicação social de Portugal. Alexandre (2000), com dados do "*Corpus* de Referência do Português Contemporâneo – oral" (CRPC), dados recolhidos a partir de programas televisivos (de notícias ou de debates) e de conversas informais com os mais variados interlocutores.

[9] *Corpus* Dialectal com Anotação Sintática - CORDIAL-SIN, coordenado por Ana Maria Martins.

7.3 Pronome relativo vs. complementador

Definir as características lexicais do introdutor de sentenças relativas é ponto central da análise aqui defendida. O objetivo desta seção é o de apresentar fatos de várias línguas que apoiam a seguinte hipótese em relação aos mecanismos de relativização nessas comunidades: a de uma especialização das formas de PRel para as relativas livres e a da forma do complementador para as relativas com antecedente (análise detalhada nas seções 7.3.2. e 7.3.3.)

7.3.1 Distinguindo os relativizadores: pronome relativo vs complementador

A evidência mais robusta da natureza nominal do PR é a manifestação de relação de concordância de traços-j (gênero, número, animacidade, etc.) entre o PR e o núcleo nominal a que correfere, além de poder apresentar flexão de caso, uma propriedade típica de DPs. Por outro lado, o complementador é uma partícula invariável, não reflete concordância nem manifesta caso. O inglês apresenta uma distribuição bem nítida entre as duas formas (exemplos de RADFORD, 1988, p.481-482):

(30) a. someone [*whom* I met —] someone [*that* I met —]
b. the book [*which* I read —] the book [*that* I read —]
c. the day [*when* we went to Paris —] the day [*that* we went to Paris —]
d. the place [*where* we stated —] the place [*that* we stated —]
e. the reason [*why* I went there —] the reason [*that* I went there —]

O alemão, por outro lado, apresenta exemplos claros de manifestação de concordância e de caso em relativas (KEENAN, 1985, p.149):

(31) German (STERN; BLEILER, 1961 apud KROEGER, 2004)

a. der Reisefuhrer, [*der* uns die Stadt zeigt]
the guide who(SG.MASC.NOM) us the city shows
(o guia que nos mostrou a cidade)

b. der Reisefuhrer, [*dessen* Adresse wir haben wollen]
the guide who(SG.MASC.GEN) address we have want
(o guia cujo endereço nós queríamos ter)

c. der Reisefuhrer, [dem ich ein gutes Trinkgeld gegeben habe]
the guide who(SG.MASC.DAT) I a good tip given have
(o guia a quem eu dei uma boa gorjeta)

d. der Reisefuhrer, [*den* ich Ihnen empfehlen kann]

the guide who (SG.MASC.ACC) I you recommend can

(o guia que eu posso recomendar a você)

Em alemão, a forma do PR é idêntica à do artigo definido, concordando com o núcleo nominal em gênero e número, além de ser marcado para caso, segundo a função gramatical que está subjacente ao núcleo nominal na SRel.[10]

As traduções dos exemplos para PB mostram que só as formas *cujo* e *quem* apresentam manifestações morfofonológicas distintas, o primeiro realizando uma forma genitiva, e o segundo, a forma dativa. Nos outros dois casos, nominativo e acusativo, a forma é sempre *que*. Isso não significa dizer que no PB os relativizadores genitivo e dativo sempre apresentam concordância de traços-j, e marcas explícitas de caso, pois os mesmos exemplos também poderiam ser realizados com o morfema *que*, uso mais difundido e mais aceito na fala e na escrita brasileiras:

(32) a. o guia *que* nós queríamos ter o endereço (dele)

b. o guia *que* eu dei uma boa gorjeta (a/pra ele)

Os exemplos em (32) e as versões em (31) mostram que o relativizador *que* pode ser usado para qualquer tipo de função relativizada, indicando mais para uma análise como complementador e não como PR. As formas com PRs nas traduções em (31b = cujo) e (31c = a quem) são ainda atestadas na escrita formal, raramente na fala, mesmo na fala planejada; assim, podem ser caracterizadas como resíduos históricos, ainda mantidos na norma padrão pela escolarização (CORRÊA, 1998).

Dizer que as formas dos PRs são resíduos históricos no PB não implica necessariamente que as duas formas não possam conviver em uma mesma língua. Em inglês, por exemplo, os dois tipos de relativizadores podem ocorrer, embora haja contextos em que o complementador está excluído.[11] Os seguintes exemplos ilustram a distribuição de PR e complementador em inglês (KROEGER, 2004, p.177):

(33) a. the spy *who* loves me

the spy *that* loves me

b. the spy *who(m)* I love

the spy *that* I love

c. the spy *from whom* I bought these documents

*the spy *from that* I bought these documents

---

[10] Em geral, o PR realiza o caso que corresponde à função relativizada. Contudo, em algumas línguas, como latim e grego, a marca de caso do pronome relativo corresponde à função do seu antecedente na sentença matriz (KROEGER, 2004).

[11] Além das diferenças morfofonológicas, evidentemente.

d. the spy *who* I bought these documents *from*

the spy *that* I bought these documents *from*

e. the spy *whose sister* I love

*the spy *that sister* I love

Os PRs do inglês são derivados de pronomes interrogativos, diferentemente do que foi visto em relação ao alemão, em que os PRs são derivados de demonstrativos. Os exemplos em (33) mostram que PRs e complementadores se distribuem livremente na realização das relativas em (33a), (33b) e (33d); contudo, a forma do complementador *that* está excluída das formas (33c) e (33e), uma vez que *that* não pode ocorrer em relativa *pied piping* (só PRs podem ser preposicionados), nem substituindo o PR *whose* (só PRs podem ser modificados nominalmente).

Os exemplos de (34) a (38) mostram que, no PB, também há restrições para a variação entre PR e complementador:

(34) a. o espião *que* me ama

b. *o espião *quem* me ama

(35) a. *o espião *quem* eu conversei

b. o espião *que* eu conversei

(36) a. o espião *de quem* comprei estes documentos

b. *o espião *de que* eu comprei estes documentos

c. o espião *que* eu comprei estes documento (na mão) dele

(37) a. *o espião quem eu comprei estes documentos *de*

b. *o espião *que* eu comprei estes documentos *de*

(38) a. o espião *cuja irmã* eu amo

b. *o espião *que irmã* eu amo

c. o espião que eu amo a irmã (dele)

d. o espião que eu amo sua irmã

A agramaticalidade de (34b) mostra a impossibilidade do PR *quem* ocorrer em função não regida por preposição, diferentemente de *who* do inglês; (35a) é agramatical por não ter a preposição realizada; ao contrário, (36b) é agramatical pela realização da preposição com o complementador. Estão excluídas as possibilidades em (37), por o PB não aceitar a preposição encalhada; como no inglês, o PB não admite o complementador *que* seguido de um elemento nominal (38b), mas pode realizar uma relativa de genitivo com o complementador *que*, em geral com um pronome possessivo anafórico na SRel (38c, d).

O quadro geral de introdutores de relativa no português (PE e PB), baseado na tradição gramatical, considera as possibilidades que constam do Quadro 1.

INVARIÁVEL	VARIÁVEL	CASO	LIVRE (SEM ANTECEDENTE)	RESTRITIVA (COM ANTECEDENTE)	APOSITIVA (COM ANTECEDENTE)
Que		-	-	+	+
Quem		-	+	+	+
Onde		OBL	+	+	+
O qual	-	-	+	+	
Cujo	GEN	-	+	+	
Quanto	-	+	+	+	
O que[12]		-	+	-	+

Quadro 1 - Distribuição dos morfemas -*wh* pelas várias estratégias de relativização em PE
Fonte: Adaptado de Alexandre, 2000, p.31.

Brito (1991) distribui o morfema *que* do PE em dois tipos: (i) nas relativas de funções não preposicionadas (sujeito e objeto relativo), a forma *que* não é um PR, ou seja, é um complementador; (ii) nas relativas preposicionadas, a forma *que* é um PR sem traços-j de concordância.

Considero que, no vernáculo dos informantes em estudo, só existe a forma *que* complementador, para qualquer um dos tipos de relativa com antecedente. Os PRs só são usados nas relativas livres. Evidências adicionais para esta proposta são apresentadas na próxima seção.

7.3.2 Tipo de relativizador: especialização de formas

A distribuição entre PR e complementador é sistemática nos dados dos informantes, os pronomes só sendo realizados nas relativas livres (exceto quanto ao *onde*; cf. seção 7.4. para uma explicação desse fato). Trato de cada caso a seguir.

7.3.2.1 O pronome QUEM

No PB, o pronome relativo *quem* é uma forma invariável quanto a gênero e número (traços-j); marcado, contudo, quanto ao traço animacidade. Nas relativas livres, não há um antecedente explícito, e o pronome *quem* realiza uma referência implícita equivalendo a *uma pessoa x*, com o traço [+animado].[13] Do total de 583 dados, há 8% de ocorrências do pronome *quem*, todas elas em estruturas de relativas

[12] Sobre as formas *os que*, *as que*, variáveis em gênero, cf. relativas semilivres, na seção 7.3.2.2.

[13] Said Ali (1921 [2001]) chama atenção para usos, na história do português, do pronome *quem* com antecedente inanimado, como no seguinte exemplo de Camões: (i) Desta vaidade *a quem* chamamos fama (Lus. 4, 95) (p. 87).

livres. O PR exerce a função sintática de sujeito, e a SRel está encaixada em diferentes posições sintáticas, como nos exemplos a seguir:

(39) a. SU: *quem* labuta com gente véia é um sofrimento doido (CZ-08)

b. OD: e eu num tenho *quem* trocasse o cheque (CZ-06)

c. OI: manda só **pá** *quem* tem dinhêro, não... (CZ-08)

d. OBL: ***com*** *quem* morava lá (HV-13)

e uma banda *de quem* trabaia. (SP-09)

porque garimpo é *quem* conhece garimpo (RC-26)

Sempre que há um antecedente explícito, a relativa se apresenta com a forma não marcada *que*, forma que não realiza qualquer tipo de traços-j, nem o traço de animacidade; ressalta-se que o pronome *quem* está excluído deste contexto.

(40) a. *O home que* mora na zona rural, (CZ-06)

b. *Pessoa que* num saiba lê é cego. (SP-09)

Esses usos são os esperados, desde que, no português, o pronome *quem* está restrito aos contextos preposicionados. Nesse tipo de contexto, a opção dos informantes do *corpus* em estudo é sempre pela relativa cortadora, com ou sem lembrete (respectivamente, (42a, b) e (41c)), todas com antecedente explícito, evidentemente:

(41) a. cê topa uma pessoa lá *que* cê trabaia dereitin *com ele* (CZ-06)

b. Mas teve um prefeito... um prefeito... aí *qu*'eu gostei *dele* (RC-26)

c. a muié de lá de Salva... de Jequié, essa Rosa, *que* eu tô falano, veio aqui olhô... (CZ-06)

A ausência de realização do pronome *quem* em relativas com antecedente, como em:

(42) a. a pessoa *para quem* dei ...

b. a pessoa *com quem* falei ...

aponta para a agramaticalidade dessa estrutura na gramática dos informantes em estudo; por outro lado, relativas 'preposicionadas' com o pronome *quem* são gramaticais na estrutura de relativa livre (cf. exemplos em (39c, d).

7.3.2.2 A forma relativa O QUE

O PR *o que* só é usado em relativas livres (3% dos dados), como ocorre em qualquer variedade do português.

(43) a. SU: agora *o que* me 'trapaia mais é uma zoadinha.... (RC-26)

O qu'eu tenho é isso! (HV-20)

b. OD: nóis reza *o que* minha mãe me ensinô, (CZ- 08)

É. Mas a gente caçô *o que*... *o que* num guardô. (RC-26)

Há algumas variações dialetais no uso de *o que* que merecem registro. Por exemplo, a forma *o que* pode ter como antecedente um nome [+humano], em lugar de *quem*, forma mais usual nas variedades urbanas do PB:

(44) a. Então, *o que* num pode ir lá, vai em Jequié (CZ-06)

b. que foi *o que* troxe casa de farinha, trôxe energia, (CZ-06)

c. E... e *o que* não fosse naquele carro aquela hora, [ele] num ia. (RC-26)

d. Não, *o que* era a moça morreu, moço! só tem eu e Romano... (RC-08)

Talvez esses sejam casos de relativas semilivres, embora sem desencadear concordância de gênero no exemplo em (44d) (sobre variação de gênero, cf. capítulo 11). Em outros exemplos, a concordância é realizada, como em:

(45) a. *As que* é feia, eu tranco o rádi... (CZ-08)

b. *A que* tava era de Bernardino, que era o mais velho. (HV-20)

c. entrega pa o… *os que* t'aí, né? (HV-20)

Há um único caso de *o que* com antecedente, o pronome indefinido *tudo*:

(46) *tudo o que* vale é o interesse e o respeito e a consideração. (CZ-06)

7.3.2.3 O adverbial ONDE

O uso deste PR é muito variável nestes dialetos. Pontuo cada tipo de ocorrência a seguir.

Primeiro, é bastante frequente nas relativas livres, com 67% dos usos de *onde*. Há 48 ocorrências do relativo *onde* (8,2% do total de relativas), sendo 32 casos de relativas livres (47a-d), algumas em contexto de fragmentos de discurso (47e-f), contextos em que não se espera variação com o marcador relativo *que*, como nos seguintes exemplos:

(47) a. mas *aonde* cê é mais acostumado é *ondé* que cê tem mais parente (CZ-06)

b. É, lá é *aonde* eu tabalho. (SP-12)

c. nóis bota, *onde* minha mãe, hoje, tirô a água e jogô fora... (CZ-08)

d. Então, *onde* a gente acha essas oportunidades, gente... a gente não vá com medo, (CZ-06)

e. *onde* tem a casa, *onde* tem sua morada (CZ-06)

f. *onde* tá parente (CZ-06)

Nas relativas livres, a preposição exigida pelo verbo matriz é realizada, como nos seguintes exemplos:

(48) a. mas vai *p'onde* foi a gente vai pr'aqui, vai pr'aculá, (CZ-06)

b. Agora se o cabra chega, não faz... não receba, não dá, não faz onde vivê, ele pode ir *pá onde* fô, ninguém presta, ninguém vale nada, né? (CZ-06)

Segundo, é o único PR usado em relativas com antecedente, ou seja, com antecedente nominal expresso (exemplos em 49), embora seja mais frequente o uso da forma *que* em estrutura de relativa cortadora (exemplos em 50):

(49) a. **Em Conquista, em Poções, na Lagoa D'Água**... *onde* eu ando aí, (CZ-06)

b. **no lugá** *donde* cê tava. Sai **do lugá** *dond'*ele tava e vai **pro lugá** *dond'*a pessoa tava. (RC-08)

(50) a. **no lugá** *qu'eu* morava, **na fazenda** *qu'*eu morava, (SP-09)

b. esses **mei canto** *que* eu rodo, eu já lhe vi, (CZ-06)

Terceiro, *onde* pode exercer funções distintas, além de localização física:

(51) a. LUGAR: luga *onde* a gente acha essas oportunidades (CZ-06)

b. OBJETO: na fila *onde* ele enfrenta, ele curte sereno (CZ-06)

c. TEMPO: então eles sai tudo pá fora e aí é *n'onde* a gente deve tê cuidado, (CZ-06)

Exerce a função de localização física em (51a); de objeto direto em (51b) (enfrentar a fila), concordando com o antecedente locativo (na fila); com valor temporal (aonde = quando), em (51c).

Quarto, o informante RC-08, com 50 anos, que já viveu fora da comunidade (seis meses em São Paulo), apresenta uma variação entre as formas *onde~donde~adonde* também presente na fala dos demais informantes, mas de maneira menos acentuada:

(52) Se ele enfezá, ele não sai, não. É... ocê tem que ir com cuidado, que no lugá *onde* ocê tá ir no lugá *adonde* ele tá, ele... quando cê tá. Cê vem com sentido *onde* ele tá e... pisa no fucinho, no lugá *donde* cê tava. Sai do lugá *dond'*ele tava e vai pro lugá *dond'*a pessoa tava. (RC-08)

Quinto, aceita retomada anafórica adverbial, seja na versão com *onde* (exemplo em 53a), seja na versão com *que* (exemplo em 53b):

(53) a. Não, per'aqui a... *aonde* a gente mora **aqui** (SP-12)

b. Essa lagoa... lugá *que* Alícia mais Ilton mora **lá**, (HV-13)

Sexto, ocorre com duplo preenchimento do sintagma complementador:

(54) a. *nonde que* nóis nasceu é bonito, né? (CZ-08)

b. *ondé que* eu planto, num sai nada... (CZ-08)

c. Aqui é *onde que* eu convivo dia-a-dia (CZ-06)

No exemplo em (54b), a forma do complementador é a mesma que se observa nas interrogativas no PB em geral: *Onde é que ....?*

7.3.2.4 QUANDO, COMO, CUJO

A forma usual das estratégias temporais é com as relativas com antecedente + *que*:

(55) a. **nessa época** *que* nós tá hoje, (CZ-06)

b. **no ano** que eu casei tava fazeno esse colégio. (CZ-06)

c. **No dia** *que* o pade Joselito faz a missa aqui (CZ-08)

d. **na hora** *que* esquenta assim tamém é... é pra chovê. (RC-26)

Só há um caso de uso do PR *quando*[14] em relativa com antecedente, construção já apresentada em (5c), um caso de relativa 'extraposta'.

(56) **naquele tempo** era bifase, *quando* eles botô, (CZ-06)

Não há qualquer caso do relativo *como*, mesmo em relativa livre. Os valores de modo do *como* são realizados pela expressão *do jeito + que*:

(57) a. *do jeitchim que* nós tinha feito... do tempo de meu pai, do meu avô...(CZ-06)

b. *Do jetcho que* eu dêxei lá, eu num vi não... (CZ-08)

Também está ausente o relativo *cujo*, como esperado. Mesmo no PE, segundo Arim, Ramilo e Freitas (2004, p.9), "o pronome cujo é cada vez menos frequente na língua falada, podendo gerar alguma estranheza quando utilizado".[15] Alguns exemplos de uso do *cujo* variando com *que*, no PE (respectivamente, exemplos (37) e (34) de Arim, Ramilo e Freitas, 2004):

(58) a. O CENELEC é um organismo de normalização europeu cujo seu domínio de actividade é dedicado exclusivamente ao sector electrotécnico. (internet)

b. Agora o Santo Padre fez uma encíclica entre a fé e a razão que eu recomendo a leitura. [*Casos de Polícia*, SIC] (cuja leitura)

mostram que o pronome *cujo* está sendo substituído pelo relativizador *que* (exemplo em 58b); quando realizado morfologicamente, como em (58a), apresenta erro de *performance* (*cujo seu*), indicando ser uma estratégia de aquisição tardia, via escolarização. Enfim, o que se pode pressupor é que *cujo* não era robusto na fala dos portugueses que aqui chegaram.[16]

7.3.3 Conclusões parciais

A hipótese defendida aqui é a de que, no processo de transmissão linguística irregular (LUCCHESI, 2000a, dentre outros), característico do contexto multilingue e multidialetal das origens da sociedade brasileira, africanos e descendentes fizeram

[14] Côrtes Junior (2006, p. 73) apresenta o seguinte exemplo, do informante 10 de Cinzento, do *corpus* base do Projeto *Vertentes*: (i) Quano batizô foi nessa 'casião (CZ-10)

[15] No *corpus* do PE estudado por Arim, Ramilo e Freitas, há 67% da forma *cujo* na escrita e 33% na fala.

[16] Cf. Jesus (2002) para um estudo do *cujo* no século XIX, no PB escrito.

reanálise morfológica e estrutural do constituinte introdutor de relativas. As reflexões são:

I. o paradigma de pronomes relativos é o mesmo dos pronomes interrogativos:

(59) a. *O que* você viu foi isto — *O que* você viu?

b. *Quem* você viu foi ela — *Quem* você viu?

c. *Onde* você viu o livro foi ali — *Onde* você viu o livro?

d. *Quando* você viu o livro foi ontem — *Quando* você viu o livro?

e. *Quanto* ele pagou pelo livro foi R$50,00 — *Quanto* ele pagou pelo livro?

II. provavelmente, no processo de aquisição irregular, as relativas livres foram assimiladas ao padrão das interrogativas; evidência clara para esta hipótese está nas realizações de relativas livres do tipo: *PR é que*, como acontece nas interrogativas (cf. seção 7.3.2.3.). O paralelo estrutural entre interrogativas e relativas livres está esquematizado abaixo, de forma simplificada:

(60) a. interrogativa indireta:

João perguntou [$_{CP}$ o que $\mathbf{C}_{+QU;+Q}$ [$_{IP}$ você viu t_{oque} ontem]]

b. interrogativa raiz: [$_{CP}$ o que $\mathbf{C}_{+QU;+Q}$ [$_{IP}$ você viu t_{oque} ontem]]?

c. relativa livre: [$_{CP}$ o que $\mathbf{C}_{+QU;-Q}$ [$_{IP}$ você viu t_{oque} ontem]] [$_{IP}$ foi $t_{CP/DP}$ isto]]

As informações necessárias para o componente fonético e para a interpretação semântica estão codificadas em C: um C_{+Q} receberá a prosódia e a interpretação de uma pergunta; um C_{-Q}, a de uma relativa livre.

III. o complementador *que* nas variedades de africanos e afrodescendentes serviu para introduzir qualquer tipo de subordinada, inclusive as relativas com antecedente, sendo assim um complementador e não um PR, como representado em (61):

(61) relativa com antecedente:

[$_{IP}$ [$_{DP}$ o lugar [$_{CP}$ OP que [$_{IP}$ você viu t_{DP} ontem]]] foi este]

IV. sendo um complementador, as relativas cortadoras foram as únicas possibilidades licenciadas, pois relativas *pied piping* requerem o uso de um pronome relativo.

A proposta acima explica a distribuição dos PRs e do complementador, exceto no que diz respeito aos usos de *onde* e *quando* em SRel com antecedente e à ausência, nos dados, dos PRs *como*, *quanto* em relativas livres. Considero que a ausência de *como* e *quanto* em relativas livres é acidental, uma questão de *corpus*. Para os usos de *onde*, assumo que este item tanto pode funcionar como um PR ou como um complementador, haja vista a possibilidade de realizar a função de objeto direto, como no exemplo em (51b);[17] nas relativas livres, o PR *onde* pode ser seguido de *é*

[17] A função sintática de *onde* no PB é bastante variável, como vários estudos apontam (CACCIAGUERRA, 2006, dentre outros).

que; nas relativas com antecedente, como complementador, não. As representações para *onde* são:

(62) a. relativa livre: $[_{CP}$ onde (é que)$_{-Q}$ ....

b. relativa com antecedente: $[_{DP}$ o lugar $[_{CP}$ OP onde$_{-Q}$ ~ que$_{-Q}$ ....

Sobre *quando* em relativa com antecedente, uso muito raro, pode ser visto como inovações mais recentes, resultantes de diferentes tipos de contato com a fala urbana. Não se pode deixar de comentar que variações observadas na fala dessas comunidades podem resultar de desenvolvimentos posteriores, oriundos do contato com outras variantes do português europeu / brasileiro, em diversas fases de constituição do PB.

Para concluir esta seção, é preciso salientar que a análise aqui esboçada não deve ser vista como uma proposta geral para a gramática de todos os africanos e afrodescendentes que aprenderam português no Brasil. As diferentes situações sociais em que viveram, os diferentes tipos de contato (mais ou menos intenso) com a língua-alvo, o acesso ou não à 'escolarização', todas estas situações devem ter produzido gramáticas diversas do português. Por exemplo, Ribeiro e Figueiredo (2006) mostram que africanos com certo domínio da escrita, no século XIX, produziam relativas *pied piping*:

(63) a. Comisçaõ Reonida, finalizaraõ o seos trabalho ***para o que famos nomiados,*** (Gregório Manuel Bahia em 29 de dezembro de 1834). (exemplo (35b) de Ribeiro e Figueiredo, 2006)

b. Comparecerem em h hum estraordinario p*ar*a o Comprim*en*to dos desvalido ***em que esta mos em Caregado*** (José Fernandes do Ó em 13 de novembro de 1836). (exemplo (35d) de Ribeiro e Figueiredo, 2006)

Desse modo, a gramática desses informantes do século XIX pode apresentar formas mais próximas da língua-alvo, do que a de informantes que adquiriram o português sem passar por algum tipo 'escolarização'. Infelizmente, a diacronia só fornece dados de escrita.

7.4 Conclusão

Uma análise superficial dos dados de língua-E, na comparação ampla entre as realizações atestadas em dialetos rurais do PE e o dialeto rural de afrodescendentes aqui em estudo, pode levar à conclusão de que a estratégia de generalização do relativizador *que* resulte de influência do PE que aqui chegou (deriva), quer dizer, os colonizadores portugueses já apresentavam em suas falas as estratégias de relativização que hoje caracterizam fundamentalmente a sintaxe dessa construção na fala brasileira. Argumento contra essa ideia nos termos a seguir.

Primeiro, as estratégias de relativização *pied piping* e *cortadora* podem estar presentes ou ausentes nas línguas humanas, de ramos e famílias independentes; é o que os estudos de Keenan e Comrie (1977, 1979) têm demonstrado amplamente. Segundo, os processos subjacentes às estratégias de relativização não são sempre os mesmos, confusão originária do uso geral de termos como relativa cortadora, relativa lembrete, relativa *pied piping* na observação superficial do fenômeno. Terceiro, usos gramaticais de tipos de relativas no PE são agramaticais no PB rural, daí sua ausência sistemática dos dados. Por exemplo, relativas *pied piping* estão presentes na fala rural portuguesa, mas totalmente ausentes na fala rural de afrodescendentes.

Segundo, em qualquer situação de aquisição de uma língua, os dados linguísticos que servem de *input* na aquisição se constituem de um subconjunto dos dados produzidos pelos adultos (LIGHTFOOT, 1991); contudo, os dados devem ser robustos para fornecer pistas para a sua aquisição. No processo de aquisição, reanálises podem ocorrer, em diferentes níveis da gramática. Na aquisição do português pelos africanos e seus descendentes, os pronomes foram reanalisados como formas específicas de relativas sem antecedente; o complementador *que* se generaliza para as formas de relativas com antecedente. Não é isso o que se observa nas variedades do PE.

Terceiro, os dados analisados também não fornecem evidências diretas para a hipótese da crioulização prévia, se se considera que as diferentes estratégias de relativização estão presentes em língua crioula, uma língua humana como outra qualquer. Alexandre e Hagemeijer (2001) estudam as SRel em diferentes crioulos atlânticos de base portuguesa,[18] atestando as seguintes possibilidades de relativização,[19] embora com distribuição diferente entre os diversos crioulos:

(64) a. *pied piping*:

Kes mudjer **ma ken** n fala e dretu. (CSV)

DEM mulher com quem 1SG falar ser simpático.

(as mulheres com quem falei são simpáticas)

b. *cortadora*:

Kes mudjer $[_{CP}$ $[_{P}\varnothing]$ **ke** n fala] e dretu. (CSV)

DEM mulher que 1SG falar ser simpático.

(as mulheres que falei são simpáticas)

[18] Crioulos da Alta Guiné e do Golfo da Guiné (da Alta Guiné: o crioulo de Santiago (CST) e o crioulo de São Vicente (CSV), falados em Cabo Verde, e o kriol (KR), falado na Guiné-Bissau; do Golfo da Guiné: o são-tomense (ST), o angolar (AN), falados na Ilha de São Tomé, e o lung'iye (LU), falado na Ilha do Príncipe.

[19] Adaptamos a terminologia aos termos assumidos no nosso texto.

c. *resuntiva*:

[Kali inen miga] ki txi ve mosu ia sa fala k'[inen] a? (PR)

que 3PL amigo KI 2SG encontrar rapaz um estar falar com-3PL PI

*'Que amigos encontraste um rapaz a falar com eles?'

d. *preposição encalhada + resuntivo neutro*:

[Inen mwala se] ku n fla **ku [ê]**. (ST)

3PL mulher DEM que 1SG falar com 3SG

(essas mulheres que falei **com ele)**

A estratégia em (64d), em que o pronome (em negrito) é uma forma invariável, sempre de 3ª pessoa do singular é estranha ao PE e ao PB; se é um indício de crioulização, está ausente nos dados do PB dos afrodescendentes em estudo.

8 As construções pseudoclivadas e clivadas

Ilza Ribeiro

Moacir Côrtes Junior

Os termos *tópico* e *foco* são usados por diferentes teóricos com diferentes significações. No sentido informal mais usual, *tópico* é o elemento acerca do qual a propriedade do predicado é afirmada, negada ou interrogada. O capítulo 9 deste livro apresenta uma análise detalhada das construções de tópico nas comunidades em estudo. Este capítulo, por outro lado, se detém na análise das estratégias de marcação de *foco*. Define-se foco em termos da noção discursiva *pressuposição*, ou seja, *o foco é a parte não pressuposta da sentença* (ZUBIZARRETA, 1998, 1999).[1] As partes pressupostas das sentenças são as assunções partilhadas pelos falantes e ouvintes quando as sentenças são pronunciadas em um dado discurso; a parte não pressuposta, ou seja, o foco, é a informação nova na sentença.

Os pares de pergunta e resposta apresentados a seguir ilustram bem essa questão:[2]

(1) O que aconteceu?

[$_F$ O gato comeu um rato].

(2) O que o gato fez?

O gato [$_F$ comeu um rato].

(3) O que o gato comeu?

O gato comeu [$_F$ um rato]

O conteúdo da pergunta realizada em cada exemplo faz parte da pressuposição; o foco é identificado como a parte da asserção que estabelece o valor do pronome interrogativo presente na pergunta. Assim, a sentença declarativa *O gato comeu um rato*, com a ordem SVO, pode funcionar como resposta às três questões acima, mas o âmbito do foco (marcado por [$_F$...]) em cada caso é diferente: toda a sentença em (1), o sintagma verbal em (2) e só o objeto em (3). Outra forma de identificar o foco, em *corpora* em que não há pares de perguntas e respostas, é através da análise do contexto pragmático-discursivo que permite reconhecer os constituintes focalizados (cf. seções 8.2 e 8.5).

---

1 Como também o definem Chomsky (1971, 1976) e Jackendoff (1972).

2 Traduzidos de Zubizarreta (1999, p.4224-4225).

Este capítulo se detém no estudo da sintaxe das estratégias de focalização identificadas como *clivagem*,[3] nos dados de fala de afrodescendentes das comunidades do *corpus* básico[4] do Projeto *Vertentes*.[5] Parte de uma descrição dos tipos de clivagem atestados no português brasileiro (PB) em geral (seção 8.1) e da caracterização discursiva do constituinte clivado como foco informacional ou identificacional, neste último caso, contrastivo ou não contrastivo (seção 8.2). Ainda na seção 8.1, formulamos algumas perguntas sobre variação e mudança na tipologia dessas construções, respondendo-as na seção 8.7. As seções 8.3 e 8.4 apresentam, respectivamente, a tipologia das pseudoclivadas (PC) e das clivadas (CLIV) atestadas no *corpus*, bem como as análises dos dados; sempre que pertinente, comparações com outras variedades dialetais do PB e do português europeu (PE) são centralizadas na discussão. As relações entre as escolhas das estratégias de clivagem segundo o estatuto informacional x identificacional do constituinte focalizado é o objeto de discussão da seção 8.5. Na seção 8.6, são apresentadas algumas conclusões sobre a sintaxe da clivagem nessas comunidades, apontando para semelhanças e diferenças em relação a outras variedades do PB. Finalmente, na seção 8.7, refletimos sobre alguns fatos diacrônicos que permitem entender a generalização da clivagem no PE e no PB, bem como a origem da *clivada sem cópula* no PB, inexistente no PE. A partir das comparações, conclui-se que a *clivagem sem cópula* resulta de uma reanálise da construção de recomplementação do PE, o que oferece indícios para a discussão dos efeitos de contato linguístico e de transmissão linguística irregular, foco central dos capítulos que compõem este livro.

8.1 Tipos de clivagem

A partir de uma distinção entre *sentenças básicas* — também denominadas sentenças neutras ou sentenças não marcadas — e *sentenças marcadas*, incluem-se as construções de *clivagem* no segundo grupo, por geralmente serem vistas como um tipo relacionado com focalização, em que um constituinte X é o *foco* identificacional, contrastivo ou não contrastivo, em estruturas do tipo *é/foi X que/quem...*[6]

[3] O termo *clivagem* será usado neste texto englobando sempre duas estratégias de focalização: pseudoclivada e clivada.

[4] Os dados coletados e as quantificações apresentadas seguem, basicamente, o estudo de Côrtes Júnior (2006).

[5] As amostras de fala analisadas no Projeto *Vertentes* são apresentadas no capítulo 5 deste livro.

[6] A tipologia de foco aqui assumida segue Kiss (1998).

Os tipos mais básicos de realização focal no PB estão exemplificados em (4b-e):

(4) a. Maria encontrou *João* — Sentença neutra

b. Foi *João* que Maria encontrou — Clivada básica

c. *João* é que encontrou Maria — Clivada invertida

d. Quem Maria encontrou foi *João* — Pseudoclivada básica

e. *João* foi quem Maria encontrou — Pseudoclivada invertida

O constituinte em itálico no exemplo em (4a) marca o foco informacional (ou foco largo), sobre o qual recai o acento nuclear da sentença; nesse caso, portanto, não há nenhuma marcação prosódica ou sintática especial sobre o foco informacional, daí o nome *sentença neutra*. Nos demais exemplos em (4), os constituintes destacados indicam a posição do acento sobre o foco identificacional, ou quantificacional, ou foco estreito.[7] É o foco quantificacional que se relaciona com as leituras semânticas do foco estreito, com leitura constante de exaustividade, mas leitura variável de contrastividade (KISS, 1995, 1998). Desse modo, os constituintes focalizados em (4b-e) identificam o valor (ou o referente) de uma variável *x*, a partir de um subconjunto exaustivo de elementos, contextualmente definidos. Em todos os quatro casos, há uma marcação prosódica especial, por isso são identificadas como *sentenças marcadas*.

Além dos tipos básicos de clivagem apresentados em (4b-e), outras possibilidades são atestadas no PB (como também em outras línguas), como as apresentadas nos estudos de Modesto (2001), e Kato et al. (1996):

(5) Pseudoclivada extraposta

a. É a Suzanita quem quer casar (MODESTO, 2001)

Pseudoclivada reduzida

b. Quero é que você vá *pra casa* (KATO et al., 1996)

Clivada sem cópula

c. *Maria* que chegou (KATO et al., 1996)

Lambrecht (2001, p.481) comenta mais um tipo de construção de focalização no PB:[8]

(6) a. Eu visitei foi ATIBAIA

b. Eu visitei Atibaia foi COM ROBERTA ontem

c. Eu visitei Atibaia com Roberta foi ONTEM

---

[7] Outros acentos secundários podem ocorrer, a depender do contexto. Observa-se que, em contexto apropriado, (4a) também pode ter uma leitura de foco estreito.

[8] Os dados em (6) são citados em Ross (1991) e retomados por Lambrecht (2001). São citados aqui a partir de Lambrecht (2001).

Esta estratégia é identificada sobretudo pela inserção da cópula antes do foco. Os exemplos evidenciam a gramaticalização da cópula como um marcador de foco, chegando a preceder qualquer argumento ou adjunto verbal para marcar a incidência do foco; mostram também que a ordem SVO é preservada nessa estrutura de focalização. Os dados em (6) se assemelham às pseudoclivadas reduzidas, motivo pelo qual foram quantificados juntos na análise do *corpus* (cf. seção 8.3).

Embora as clivadas sejam construções que ocorrem em muitas línguas, a distribuição dos seus tipos não é uniforme entre as línguas. Segundo Moreno Cabrera (1999, p.4281), não há construções clivadas no espanhol peninsular moderno; entretanto, todos os dialetos da América desenvolveram estratégias de clivagem. Em suas palavras: "En los dialectos de América está muy extendido un tipo de perífrasis en la que en vez de la relativa libre aparece una oración o sintagma verbal encabezado por la conjunción *que*".[9] São exemplos (itálico acrescentado):

(7) a. Fue *en el siglo XV* que se descubrió América.

b. *En una escalera* fue que reñimos.

c. Fue *aquí* que le dio el ataque.

d. Es *el barbero* que lo rasura.

Salienta o autor que há exemplos desse tipo de clivada (perífrase conjuntiva) na tradição literária espanhola, em Lope de Vega (1562-1635) e Quevedo (1580-1645):

(8) a. *Es* por celos de su amor *que* injustamente le afrenta (F. Lope de Vega, *D. Juan de Castro*, 2ª. Parte, acto I, esc. VI)

b. Por vos *es que* yo he sufrido el oprobio, y que la confusión há cubierto mi rostro.

O estudo de Moreno Cabrera (1999) aponta para dois fatos diacrônicos importantes: (i) o espanhol que chega na América realiza construções clivadas, ao menos de constituintes não argumentais; (ii) o espanhol da América amplia o uso dessas construções, enquanto se tornam obsoletas no espanhol peninsular; (iii) as pseudoclivadas são possíveis nas duas variantes dialetais. Construções como as em (8) só podem ser realizadas como pseudoclivadas no espanhol peninsular moderno, mas podem ser clivadas no espanhol americano.

As diversas possibilidades de clivagem nos dados sincrônicos do PB e as divergências de possibilidades entre pseudoclivada e clivada no espanhol, peninsular e americano, levanta três perguntas cruciais: (i) o que distingue as variedades

[9] Moreno Cabrera (1999) utiliza uma terminologia diferente da dos estudos da gramática gerativa: perífrase de relativo para as pseudoclivadas; e perífrase conjuntiva, para as clivadas. Mantemos, na apresentação, a terminologia da gramática gerativa.

portuguesa e espanhola americanas da variedade espanhola peninsular? (ii) as duas variedades do português (PE e PB) também apresentam uma distribuição diferenciada entre as duas possibilidades de clivagem? (iii) quais evidências os dados das comunidades de afrodescendentes fornecem para responder às questões em (i) e (ii)? Retomamos este tópico na seção 8.7, quando argumentamos que as diferenças essenciais entre dialetos que permitem as construções clivadas e dialetos que as recusam se fundamentam na gramaticalização do verbo copulativo *ser* e o complementador *que* em uma forma gramatical especializada para marcar foco: *é ... que*.

8.2 Pseudoclivadas e clivadas: estrutura e estatuto informacional

Como já comentado acima, as *pseudoclivadas* (PC) e as *clivadas* (CLIV) são construções que permitem realçar uma informação nova, ou foco da sentença, ou seja, as construções de clivagem são usadas para codificar sintaticamente o fenômeno de nível discursivo conhecido como foco. Vimos que os contextos de perguntas e respostas são os mais evidentes para se observar a clivagem, o que ilustramos agora com dados do *corpus* em estudo:

(9) a. DOC - A senhora nunca foi a São Paulo?

INF - **Quem** sempre vai a São Paulo **é** MEU MARIDO. (RC-13)

b. DOC: A senhora foi no cozinhado?

INF: Ieu num fui não... **quem** foi **foi** INÊS. (CZ-10)

(10) a. DOC: Quem deu chapéu? Comprô pra isso ou...

INF: O chapéu que eu usei mesmo ***foi*** *Ana* ***que*** me deu (CZ-01)

b. DOC: E quem fundô isso aqui?

INF: Ah, (**foi**) *esses mais véio* (**foi**) **que** fundô isso aqui. (CZ-11)

Os exemplos em (9) são de PCs; em (9a), o constituinte MEU MARIDO é um foco contrastivo,[10] pois, na resposta da informante, há um contraste entre ela e o seu marido, no que diz respeito à ida a São Paulo. Em (9b) o foco INÊS também é contrastivo, fazendo oposição entre Inês e a própria informante.

A PC é uma construção copulativa em que o verbo copulativo SER seleciona uma minioração cujo sujeito é uma relativa livre (cf. capítulo 7) e o predicado, o constituinte focalizado:

[10] Os constituintes focalizados estão sempre em itálico; quando se quiser realçar o valor contrativo do foco, este será representado em caixa alta.

(11) SER_{cop} $[_{SC}$ Relativa livre foco $]$[11]

Há mais de uma possibilidade de ordenação superficial dos dois constituintes da minioração, o que está na origem dos diferentes tipos de PC, tema que tratamos nas subseções de 8.3, a seguir.

O contexto em (10a) exemplifica o uso da CLIV, em resposta a um pedido de informação. Nesse caso, o foco é identificacional, e não contrastivo, embora uma clivada também possa ser usada para codificar foco contrastivo. As duas possibilidades de realização da cópula em (10b), sinalizadas entre parênteses, apontam para dois fatos: (i) o constituinte focalizado pode estar em posição pré- ou pós-cópula; (ii) a cópula pode estar ausente. Na seção 8.4, detalhamos as propriedades das clivadas no *corpus*.

A estrutura da clivada é também a de uma copulativa, do seguinte tipo:

(12) SER_{cop} $[_{CP}$ que $[_{TP}$ ...... Foco .....]

Na representação em (12), o verbo *ser* seleciona um complemento CP, e não uma minioração, como na PC. Na realização superficial da CLIV, o constituinte focalizado –– sujeito, objeto ou adjunto –– se desloca de TP para a posição apropriada de foco identificacional, contrastivo ou não. Como há mais de uma posição disponível para o pouso do foco, também há mais uma ordenação superficial de constituintes, o que resultará em diferentes estratégias de clivagem, discutidas nas subseções de 8.4.

8.3 Tipos de pseudoclivadas nas comunidades em estudo

Nesta seção apresentamos os diferentes tipos de PC atestados nas comunidades em estudo, com as suas principais características sintático-discursivas. Foram computados 376 dados de clivagem, sendo 30,5% de PCs.

8.3.1 Pseudoclivada básica

Denomina-se *pseudoclivada básica* a PC que se realiza superficialmente com o foco realizado pós-cópula, ocorrendo em 25% dos dados.

(13) $[_{TP}$ RLivre cópula $[_{VP}$ t_{cop} $[_{SC}$ t_{RLivre} **foco**]

[11] Todas as representações estruturais serão apresentadas de forma simplificada, pois não é objetivo central deste texto desenvolver uma análise formal da clivagem. Ao leitor interessado neste ponto, indicamos a leitura de: Kato e Ribeiro (2007), Modesto (2001), dentre outros citados nas Referências.

Nos exemplos em (14), o foco é identificacional não contrastivo:

(14) a. que, quem ensina no Giló é *uma muié de fora lá de Rio de Conta*, né. (RC-26)

b. É. O qu'eu tenho é *isso*! (HV-20)

c. onde nós fomo *foi pra Boa Lembrança*. (HV-20)

A PC básica também realiza o foco contrastivo, em aproximadamente 10% dos dados coletados, como em:

(15) a. poque quem profura os poço estesiamo é *a CERPE*... (CZ-06)

b. Quem custura ali era EU. (CZ-08)

c. quem fica é ELE. (CZ-11)

d. quem faz nós é NÓS MESMO. (CZ-6)

e. quem foi foi CLÁUDIA. (CZ-11)

Os dados mostram que o foco realizado em posição pós-cópula pode ser identificacional [+/– contrastivo].

8.3.2 Pseudoclivada invertida

Como o nome já sugere, o foco na PC invertida se realiza em posição invertida à de realização do foco na PC básica, ou seja, o foco se posiciona antes da cópula, como no seguinte esquema:

(16) $[_{TP}$ **Foco** cópula $[_{VP}$ t_{cop} $[_{SC}$ Relativa livre t_{foco}]

A PC invertida é de uso pouco frequente nas comunidades em estudo (1%), só tendo sido atestados dados com relativas introduzidas por *onde*, sendo o foco um pronome dêitico:

(17) a. *Aqui* **é** onde que... eu convivo dia-a-dia (CZ-06)

b. *lá* **é** aonde eu tabalho. (SP-12)

c. *aí* **é** n'onde a gente deve tê cuidado (CZ-06)

8.3.3 Pseudoclivada reduzida

A característica da PC reduzida (6% dos dados) é justamente a omissão do pronome relativo, como nos seguintes exemplos:

(18) a. eu não tenho queda **é** *pa gado*! (HV-20)

b. condo a gente tá lá vê **é** *os tiro ali naquela*... num tem uma vendinha aí? (SP-09)

c. a ladêra só faz deslizá [...] quano tá choveno muito, faz **é** *afundá*, terra de arêa. (CZ-06)

d. Agora, aqui faz **é** *da cana mesmo*. (RC-26)

e. eu compro a carne de boi **é** *ali embáxo mermo*, é ali. (SP-09)

f. mas eu compro mais **é** *da ôta*, de parafina. (SP-09)

h. ieu gosto **é** *de roça*, num gosto de comércio não. (HV-13)

Todos os exemplos em (18) poderiam ter um pronome relativo.[12] Por exemplo, (18)a. poderia ser realizada como:

(19) *Para o que* não tenho queda **é** *pa gado*

Também foram encontrados casos de PC-red focalizando sujeito, estratégia que, segundo Brito e Duarte (2003), não é permitida no PE culto:

(20) *Leu o livro foi João PE (BRITO; DUARTE, 2003, p.693)

(21) a. compô o lugá **foi** *Ségio*. (CZ-10)

b. tá vindo **é** *o padre Juselito*. (CZ-09)

Segundo Franco (2007, p.61),[13] as PC-red acontecem no PE rural com clivagem de todas as funções, exceto a de sujeito:

(22) a. Mas ele queria **era** AS TERRAS. (COV20-N)

b. tudo gosta **é** *de carne fresca.* (PST09-N)

d. Agora peguei **foi** *por curiosidade*. (MIG49-N)

e. Porque eu se tivesse lugar, eu botava **era** *aqui*. (MIG56-N)

8.3.4 Outros casos de focalização do tipo pseudoclivada

Há também algumas construções de focalização que incluímos como PC básica, embora se realizem em relativas com cabeça:

(23) a. Bom, o único que a gente pode perdê **é** *a mandioca*, (CZ-06)

b. o mais, mais faço que tinha **era... era** *fazê um saco de farinha e vendê por dez real* (CZ-06)

c. a segunda pessoa do governo aqui **é** *ele mesmo* (CZ-06)

d. a primêra associação que tá entrano no projeto aqui no semiárido da caatinga **é** *do Cinzento*, (CZ-06)

---

[12] Isso pode levar a admitir que a PC-red tem a mesma estrutura da PC básica, diferindo na estrutura interna da sentença relativa.

[13] Os dados de Franco (2007) foram coletados do *Corpus* Dialectal com Anotação Sintática (CORDIAL-SIN) coordenado por Ana Maria Martins, Universidade Clássica de Lisboa.

e. uma das coisa mais necessára que nós precisa **é** *a casa de farinha* (CZ-06)

f. o mês que deságua mesmo **é** *novembo, dezembo, janêro, feverêro e maço e abril* (CZ-06)

São todas construções de focalização, que podem também ser incluídas na classificação de copulativa especificacional, ou PC especificacional.[14] Contudo, optamos por analisar como PC básica, pois os únicos elementos que aparecem como cabeça são expressões de lista: *o único que, a primeira que, uma das que...*[15]

Casos como esses são atestados no inglês e analisados como PC especificacional. Segundo den Dikken (2001), não há uma distribuição equilibrada dos pronomes relativos em PC; a forma mais frequente no inglês é com *what*; para os demais tipos, uma relativa com cabeça é mais usual do que uma relativa livre (exemplos de den Dikken, 2001, p.9):

(24) a. *the things* John does not eat are food for the dog (what)

b. *the place* where John finally ended up was in Berkeley (where)

c. *the time* at which John arrived was at five o'clock (which)

d. *the reason* why John went to the bookstore was to buy a book about pseudoclefts (why)

e. *the way* John did it was by using a decoder (how)

f. *the {person}* who John visited was Bill (who)

Comportamento semelhante se observa nessas comunidades, em que os únicos pronomes atestados em relativas livres são: *quem, onde, o que* e *quando* (cf. capítulo 7).

Só há um caso codificado como de PC extraposta, com uma relativa com cabeça:

(25) a. **É** *essa aqui* a cana que eu falei, que eu ia mode brigá mais os home, (CZ-08)

Essa PC se caracteriza por apresentar uma ordem em que a relativa parece ter sido extraposta para o final da sentença, à direita, a partir de uma PC básica:

(26) $[_{TP}$ $t_{RLivre2}$ cópula $[_{VP}$ t_{cop} $[_{SC}$ $t_{RLivre1}$ **foco**] RLivre

Essa estrutura de focalização é atestada na diacronia do PE, sempre com uma relativa com cabeça no português arcaico (27a-c); com um pronome relativo, a partir do século XVIII (27d-e):[16]

[14] Do tipo: A causa da briga foi *você*, em que se tem uma sentença copulativa com o foco realizado como predicativo.

[15] Além disso, como apontado no capítulo 7, há uma tendência nesta comunidade a realizar relativas com cabeça.

[16] Dados de ALORNA, Marquesa de. *Inéditos - Cartas e outros escritos*. (1750-1839) Selecção, prefácio e notas de Hernâni Cidade. Lisboa: Sá da Costa, 1941. <www.ime.usp.br/~tycho/corpus>.

(27) a. Nõ **he** *amigo* aquelle que em todallas cousas nõ ama prol do seu amigo. (*Crónica Geral de Espanha de 1344*)[17]

b. ca nom **é** *este* o primeiro *torto* que el fez (*A Demanda do Santo Graal*)[18]

c. Non **é** *esta* a primeira vilania que el de mim disse (*A Demanda do Santo Graal*)

d. ***foi*** Vossa Eminência *quem* me conduziu à presença de Sua Alteza Real

e. Não ***é*** Espanha *quem* deve estabelecer as regras da nossa conduta, mas *é* o Nosso Príncipe *quem* deve ditar à Espanha o que convém;

Em outras variedades dialetais do PB, qualquer tipo de pronome relativo pode ocorrer em relativas 'extrapostas', como nos seguintes exemplos:

(28) a. É *nesta sala* **onde / o lugar onde** eu estudo

b. É *Maria* **quem / a pessoa que** faz bolos

c. É *este livro* **o que / a coisa que** eu estou procurando.

d. É *hoje* **quando**[19] / **o dia que** eu devo concluir meu trabalho.

e. É *assim* **como / o modo como** eu faço este trabalho.

8.4 Tipos de clivadas nos dados das comunidades em estudo

Foram identificados três tipos de clivadas, no *corpus*: (i) clivada básica (CLIV); (ii) clivada invertida (CLIV-inv); (iii) clivada sem cópula (CLIV-sem-cop). A estrutura básica da clivada foi apresentada em (12) acima, quando comentamos haver mais de uma possibilidade de pouso para o constituinte focalizado; são essas diferentes possibilidades que estão subjacentes aos três tipos identificados, conforme análise a seguir.

8.4.1 Clivadas básicas

Na clivada básica, o constituinte focalizado ocupa uma posição de foco interna à sentença, como representado simplificadamente em (29):

(29) $[_{TP}$ cópula $[_{VP}$ t_{cop} $[_{FocoP}$ [FOCO] [*que* $[_{TP}$ t_{FOCO}]]]]]

[17] Edição de L. F. L. Cintra. 4v. Lisboa: Academia Portuguesa da História, 1951.

[18] Edição de J-M. Piel, concluída por Irene Freire Nunes. Lisboa: Imprensa Nacional-Casa da Moeda, 1988.

[19] O uso do relativo *quando* é aceito pelos gramáticos normativos, mas muito pouco atestado na fala brasileira.

A representação realça dois movimentos sintáticos, o do verbo copulativo para o núcleo T, onde realizará os traços morfológicos de tempo e estabelecerá concordância com o foco; o do constituinte que recebe o acento focal (seja sujeito, complemento ou adjunto) para o especificador de FP, na periferia esquerda interna da sentença.[20]

Os exemplos seguintes são de *clivadas básicas*:

(30) a. a energia **é** *a Coelba* **que** pega a verba (CZ-06)

b. **É**... *lá* **que**... vai no banco recebê dinhêro... (HV-13)

c. É, **é** *minha fia* **que** mora ali. (SP-09)

d. Não, **é**... *ela* **que** toma conta, de a roça a ela, (CZ-08)

e. **É** *Jeane* **que** é noiva de um rapaz de Salvador (SP-09)

f. **foi** *um bando de filho* **que** ele teve. (SP-12)

g. **é** *dessas* **que** gente bebe aí. (RC-26)

h. **é** *Salviano* **que** conversa. (CZ-09)

Os dados em (30) mostram a focalização de diferentes tipos de constituintes do sintagma verbal: o sujeito (a, c, d, e, h), o objeto direto (f), o objeto direto partitivo (g); o adjunto adverbial (b). Contudo, a função discursiva de foco recai mais frequentemente no constituinte que exerce a função sintática de objeto (peso relativo .63), seguida de adjunto (.51); por último, o sujeito (.46).

8.4.2 Clivadas invertidas

As clivadas invertidas apresentam uma ordem em que o foco antecede a cópula; a cópula e o complementador *que* ocorrem superficialmente adjacentes, como nos seguintes exemplos:

(31) a. *Isso* **é que** eu falo. (HV-19)

b. *Isso* **é que** é ruim. (HV-20)

c. A *muié* **é que** sabe. (CZ-06)

d. e *ela* **é que** comanda o escritório da associação, (CZ-06)

e. Agora, *esses d'agora* **é que** tem uma sorte... (HV-13)

f. então *só nos pé da serra* **é que** tem água corrente, água sargada. (CZ-06)

g. *Eu* **é que** sô o valuadô de... (HV-20)

h. *Por isso* **é que** ocês morre ligêro (RC-08)

[20] Sobre FocoP interno, cf. Belletti (2004, 2005).

Constituintes com diferentes funções sintáticas podem ser focalizados por esta estratégia: sujeito (exemplos em (b, c, d, e, g) com peso relativo de .47 de probabilidade de uso); complementos verbais (exemplo em (a) com peso relativo de .30); adjuntos adverbais (exemplos em (f, h) com peso relativo de .62).

A diferença estrutural entre a clivada básica e a clivada invertida pode ser simplificadamente esquematizada como em:

(32) $[_{FocoP}$ [FOCO] $[_{F'}$ cópula $[_{TP}$ t_{cop} $[_{VP}$ t_{cop} $[_{CP}$ *que* $[_{TP}$ t_{FOCO}]]]]]

Nessa estrutura, o foco ocupa a posição de especificador de FocoP na periferia esquerda externa da sentença (não na periferia interna da sentença, como na clivada básica); em geral, são focos contrastivos que são atraídos para essa posição, mas foco não contrastivo também pode ocupar essa posição, como nos exemplos em (31 b, d) acima.

8.4.3 Clivadas sem cópula

O apagamento da cópula na construção de clivagem é bastante comum nessas comunidades (22%). Exemplos de *clivadas sem cópula* estão em (33) abaixo:

(33) a. *Eu* **que** le... *eu* **que** levei fama aquele dia. (HV-20)

b. *Ela* **que** sabe. (SP-09)

c. *ano passado* **que** ele levô nóis. (CZ-08)

d. *a mãe* **que** criô (CZ-06)

e. *Finada minha mãe* **que** criô. (HV-13)

f. *só na rua* **que** eu bebo. (RC-08)

g. *Marta* **que** chegô aqui me preguntano o que é que... (CZ-08)

h. *Muita água* **que** tem aqui, mesmo! (RC-08)

i. *Eu mermo* **que** torrava farinha... (SP-12)

Para os exemplos em (33), podemos pensar que a posição da cópula pode estar antes do constituinte focalizado, no início da sentença, o que a identificaria como uma clivada básica, ou após o constituinte focalizado, constituindo-se numa clivada invertida.

A estratégia de *clivada sem cópula* está ausente do PE culto (BRITO; DUARTE, 2003) e do PE rural (FRANCO, 2007); contudo, é uma estratégia comum no PB em geral (BRAGA, 1991; KATO et al., 1996).

Franco (2007, p.57-58) observou um grupo de construções do PE rural que poderia estar na origem das clivadas sem cópula. Contudo, são construções de recomplementação, em que um constituinte, geralmente mas não exclusivamente exercendo a função de sujeito, aparece ensanduichado entre dois complementadores:

(34) a. Parece **que** *eu* **que** ouvia falar nessa coisa. (AAL18-N)

b. Metia-se-me cá na ideia **que** *elas* **que** não eram capazes de aprender. (MST12-N)

c. Eu sei **que** *ele* **que** anda nas Rasas. (MST02-N)

d. Parece **que** *na Suiça* **que** dão muita importância a essas coisas. (AAL04-N)

e. agora ouviamos falar **que** *em Lisboa* **que** as varinas apregoavam lá carapau e aqui era charro. (PAL09-N)

Este tipo de recomplementação é atestado na história do PE desde sua fase arcaica (RIBEIRO, 1995):

(35) a. e o abade San Beento dizendo o contrairo **que** *Deus* **que** o fezera por el (*DSG*, 2.7.9[21] - século XIV)

b. mandou-lhi dizer **que** *el* **que** o ia ver (*DSG*, 2.14.4 - século XIV)

Nos documentos mais formais dos séculos XVI e XVII, analisados por Ribeiro (2007), não foi observada nenhuma ocorrência dessa construção. Contudo, em textos escritos por mãos inábeis, editados por Marquilhas (1997), seu uso é bastante recorrente. Alguns exemplos são como:

(36) a. he homem q. migou na natureza da sua mula dizendo **q**. *a mula* ***q***. estaua com dezeios de fazer tal couza (MARQUILHAS, 1997; Anexos III, Documento IV – 1617-1620) (leitura: É homem que mijou na natureza da sua mula, dizendo que a mula que estava com desejos de fazer tal cousa)

b. porq. diz **q**. *os trusquos* **q**. dizem q. a snar. q. não pari sem uarão (MARQUILHAS, 1997; Anexos III, Documento IV – 1617-1620) (leitura: porque diz que os turcos que dizem que a Senhora que não pariu sem varão)

Por outro lado, para Côrtes Junior (2006), a clivada sem cópula é uma estratégia inovadora, no que diz respeito ao PB, resultado do processo de transmissão linguística irregular (LUCCHESI, 2003). Voltamos a falar dessas duas hipóteses na conclusão, quando vamos argumentar a favor da hipótese de Franco (2007).

8.4.4 Clivadas com dupla cópula

Há alguns casos de clivadas em que há duplo marcador de foco, quer dizer, dupla realização da cópula:

(37) a. **é** *por isso* **é que** eu tem... (CZ-06)

b. aí **é** *aonde* **é que** eu tô lhe dizeno, (CZ-06)

[21] MATTOS e SILVA, R. V. *A mais antiga versão portuguesa dos "Quatro livros dos Diálogos de São Gregório". Edição crítica*. 1971. 4v. Tese (Doutorado) - Universidade de São Paulo, São Paulo.

Em estudo sobre o funcionamento das sentenças clivadas no discurso semicoloquial oral falado do Rio de Janeiro, Braga (1991) também atesta clivadas com dupla cópula, identificando-as como de "duplo foco":

(38) F: [...] O rombo tá muito grande (X)... O rombo é muito grande. Então pa eles tentar fechar, quem sofre é a gente, rapaz! **É** *a gente* **é que** sofre.

Franco (2007, p.56) também cita exemplos de clivada com dupla cópula, nos dados do PE rural:[22]

(39) a. **Era** *broa* **é que** a gente fazia! (MIN19-N)

b. **É** *só o mestre* **é que** manda. (VPA03-N)

c. **É** *só a irmã* **é que** está aí. (COV09-N)

d. **Foi** *quando ele apareceu* **é que** aconteceu isso. (COV25-N)

e. **É** *com aquele material de vido* **é que** ele faz o ninho. (CTL21-N)

Essas construções indicam claramente a gramaticalização de *é que*, como um único constituinte gramatical. Voltamos a essa questão na conclusão.

8.4.5 Outros casos de clivagem

Existem outras estratégias de clivagem, com estruturas diferentes das já tradicionalmente conhecidas por usar uma *cópula* e o complementador *que* ou a forma gramaticalizada *é que*.[23] Por exemplo, foram encontradas possibilidades de focalização com a forma verbal *ter*, em estratégias semelhantes às de relativas existenciais (cf. capítulo 7):[24]

(40) a. DOC : E vende?

INF: Vende.**Tem** *gente aí* **que** vende. (SP-04)

DOC: Ah, essa é nova, então!

INF: É... **tem** *uns seis ano* **que** fez. (SP-01)

Um outro tipo de clivagem discutido no estudo de Lambrecht (2001) é o denominado *all-cleft*. Os dois exemplos em (41a-b) exemplificam essas construções em inglês e no português de afrodescendentes, respectivamente:

(41) a. **All** this machine does **is** SWIM AND EAT.

(*Tudo que essa máquina faz é nadar e comer*)

---

[22] Nenhuma informação sobre esta estratégia no PE culto é apresentada no estudo de Brito e Duarte (2003).

[23] Cf. Lambrecht (2001) para uma apresentação sobre outras possibilidades, em diferentes línguas.

[24] Talvez semelhantes ao tipo com *have* no inglês, segundo apresentado por Lambrecht (2001).

b. **Tudo que** a gente fala **é** *graças a Deus*. (RC-13)

A diferença do tipo em (41) em relação à PC é exatamente o que lhe atribui o nome, ter o pronome indefinido *all/tudo* introduzindo a relativa.

8.5 Resultados do estatuto discursivo do foco nas comunidades em estudo

Como apontado na seção 8.2, o constituinte focalizado pode realizar um foco identificacional, contrastivo ou não. A análise dos dois tipos considerou não só os contextos de perguntas e respostas (exemplificados em (9) e (10) acima), como também a situação discursiva em que a clivagem acontecia. Por exemplo, no contexto em (42),

(42) INF.: Batida de limão, batida de gingibe.

DOC.: E a senhora sabe fazê?

INF.: Eu num sei não... eu num se... eu nunca fiz.

DOC.: Nenhum tipo de bebida?

INF.: Agora, quem faz é ali, ó, ININT pere qui tudo... pere qui todo desse municipi só num fai eu, qu'eu num... eu num tinha paciênça... eu... eu num tinha paciênça de fazê. (SP-09)

observa-se que o dêitico *ali*[25] contrasta com as informações anteriores e posteriores: *Eu num sei não... eu num se... eu nunca fiz.*[...]. *pere qui todo desse municipi só num fai eu*.

Por outro lado, o contexto de ocorrência do exemplo em (43) aponta para um foco identificacional não contrastivo:

(43) DOC: E o senhô deve ficá satisfeito, que tem até um professô dend'e casa, né?

INF: Um...pois é. Antigamente não tinha não. [...] Tem uma no Bananá tamém até prima nossa tamém, até prima minha tamém no Bananá aqui embaxo e nada disso tinha.

DOC: É o progresso, né?

INF: Pois é. E por aqui tudo já tem. No Mato Grosso... Mato Gosso... Giló... é qu'eu num sei se o Giló tem. Que, **quem ensina no Giló é uma muié de fora lá de Rio de Conta**, né. (RC-26)

O tema da conversa é a existência de professor nas comunidades; o informante, concluindo a relação das comunidades, identifica a professora de Giló: *quem ensina no Giló é uma muié de fora lá de Rio de Conta*.

---

[25] Provavelmente a informante aponta para algum participante da entrevista.

Às vezes, a interpretação entre foco identificacional contrastivo e foco identificacional não contrastivo não é muito evidente, como no seguinte contexto do exemplo em (44):

(44) DOC: Só veio um engenhêro?

INF: Aqui vêi dois, vêi um aí essa vez e vêi ôtro lá, né?

DOC: Não vieram junto...

INF: Não, vêi da CERPE, poque **quem profura os poço estesiamo é a CERPE**... a gente faz o... o projeto, o pedido e aí quando o pedido vem, aí cê vai tê que assiná o projeto e... o presidento do conselho, que é o prefeito torna a levá o governo a carta. (CZ-06)

Os falantes estão conversando sobre furar poços artesianos, vinda de engenheiros, etc.; em nenhum momento anterior a CERPE foi citada. As falas do documentador parecem indicar que ele pensa que são os habitantes do povoado que chamam os engenheiros para furar os poços. O informante CZ-06 procura negar a pressuposição, dizendo que *quem profura os poço estesiamo é a CERPE*, que os engenheiros são mandados pela CERPE, que eles só fazem o projeto solicitando a perfuração do poço. Se a leitura está correta, a PC desse exemplo introduz um foco contrastivo, negando um pressuposto do ouvinte.

Cruzando estes dois tipos de contexto, o das perguntas / respostas e o das situações discursivas, para os tipos de clivagem mais atestados, Côrtes Junior (2006) chegou aos resultados percentuais de usos do foco identificacional contrastivo vs. não contrastivo que constam da Tabela 1.

Tabela 1 - Tipo de estatuto da informação do constituinte focalizado nas sentenças clivadas

TIPO DE CLIVADA	IDENTIFICACIONAL NÃO CONTRASTIVO			IDENTIFICACIONAL CONTRASTIVO		
	Nº ocorr.	%	P.r.	Nº ocorr.	%	P.r.
CLIV	**89/123**	**37/33**	**.55**	36/123	26/33	.41
CLIV-sem-cóp	**60/81**	**26/22**	**.57**	21/81	15/22	.38
CLIV-inv	14/57	6/15	.33	**43/57**	**30/15**	**.77**
PC	57/93	24/25	.51	36/93	26/25	.48
PC-red	**17/2**	**27/6**	**.57**	5/22	4/6	.38
Total	235/37	6100	- -	141/376	100	- -

Fonte: Côrtes Junior, 2006, p.87, com adaptações.

As percentagens e pesos relativos apresentados na Tabela 1 permitem as seguintes conclusões:

(i) a *clivada básica* e a *clivada sem cópula* são favorecidas quando o foco é não contrastivo, apresentando sempre peso relativo superior (.55 e .57, respectivamente) quando comparado ao das construções com foco contrastivo;

(ii) a *clivada invertida* revela um comportamento diferente dos dois outros tipos, como atesta o peso relativo .77 a favor do foco contrastivo;

(iii) a *pseudoclivada* mostrou-se indiferente ao tipo de estatuto discursivo do foco, com o peso relativo de .51 a favor do não contrastivo e .48 a favor do contrastivo;

(iv) a *pseudoclivada reduzida* realiza-se mais com o foco não contrastivo, com peso relativo de .57.

8.6 Conclusões parciais

A quantificação das estratégias de clivagem mais usadas nas comunidades estudadas está representada na Figura 1.

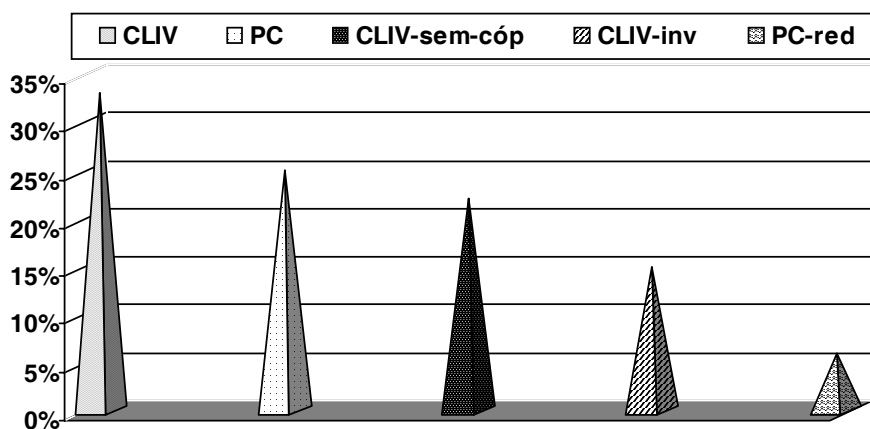



Figura 1 - Ocorrências dos tipos de clivagem nas comunidades estudadas
Fonte: Côrtes Junior, 2006, p.80.

A Figura 1 mostra uma linha decrescente contínua nas realizações das estratégias, da mais frequente (clivada básica, 33%) à menos usual (pseudoclivada reduzida, 6%). Todos os tipos permitem a focalização de constituintes de funções sintáticas diversas, como representado na Tabela 2.

Tabela 2 - Cruzamento do estatuto focal com o tipo de estatuto sintático do constituinte focalizado em cada tipo de sentença clivada

TIPO DE CLIVADA	SUJEITO		OBJETO DIRETO		ADJUNTO ADVERBIAL	
	~Cont. Nº oc./%	Cont Nº. oc/ %	~Cont. Nº oc./ %	Cont. Nº. oc/ %	~Cont. Nº oc./ %	Cont. Nº. oc/ %
CLIV	**37 / 35**	19 / 21	**17 /47**	5 / 38	**33 /35**	12 / 32
CLIV-s-cóp	28 / 26	14 / 16	3 / 8	2 / 15	29 / 31	5 / 13
CLIV-inv	6 / 6	24 / 27	2 / 6	1 / 8	6 / 6	**18 / 47**
PC	33 / 31	**32 / 36**	6 / 17	3 / 23	18 / 1	91 / 3
PC-red	2 / 2	1 / 1	8 / 2	22 / 5	7 / 8	2 / 5
Total	106 / 100	90 / 100	36 / 100	13 / 100	93 / 100	38 / 100

Fonte: Côrtes Junior, 2006, p.92.

Observando-se os resultados da Tabela 2, constata-se que a focalização de sujeito com estatuto não contrastivo favorece a CLIV com a frequência de 35%; a focalização de sujeito com sentido contrastivo favorece a PC, com 36%; na função de objeto, a CLIV apresenta as maiores frequências, 38% com sentido contrastivo e 47% com sentido não contrastivo, resultado que reafirma a preferência, nas CLIVs, pelo estatuto não contrastivo e pela focalização de objeto; quanto ao adjunto adverbial, o foco não contrastivo favorece mais uma vez a CLIV, e o foco contrastivo, a CLIV-inv.

Segundo o comportamento das sentenças clivadas aqui apresentado, quando o falante focaliza um objeto prefere as CLIVs. Pode-se supor que seja pelo fato de essa estratégia possibilitar que o falante mantenha o uso do complementador *que*, em lugar de um pronome relativo, com qualquer que seja o tipo de constituinte clivado. Nos exemplos em (45):

(45) a. **é** *minha fia* **que** mora ali. (SP-9)

b. **É** *o começo do samba* **que** eu falo po senhô. (HV-12)

c. **Foi** *na casa de dona Francilina* **que** eu tava trabaia... (RC-4)

observamos que se mantém a mesma estrutura da CLIV (*ser* + *que*), mesmo tendo a focalização de *pessoa* (45a), de *coisa* (45b) e de *lugar* (45c), em que o falante poderia usar os pronomes relativos *quem*, *o que*, *onde*, atestados nas comunidades estudadas.

Se o falante optasse por usar uma PC, em lugar de uma CLIV, necessariamente teria que usar um pronome relativo concordando com o tipo de constituinte focalizado. Isto é, se o constituinte fosse um objeto ou coisa, usaria o pronome *o que*; se pessoa, o pronome *quem*; no caso de adjunto adverbial, um pronome que concordasse com o tipo de adjunto: de modo, *como*; de lugar, *onde*; de tempo, *quando*, etc. Isso pode explicar a diminuição das ocorrências de PC focalizando adjunto, ou seja, a esquiva em usar o pronome relativo para funções adverbiais (cf. capítulo 7 sobre as restrições quanto ao uso de pronomes relativos).

Contudo, no que diz respeito às funções adverbiais, a expectativa era de um uso maior da PC-red, pois, embora seja considerada um subtipo da PC, sua estrutura permite a omissão do pronome relativo. No entanto, no geral há poucas ocorrências de PC-reduzida (6%).

O estudo de Braga (1991), sobre a clivagem no português semicoloquial do Rio de Janeiro, revela que as CLIVs focalizam tanto sujeito e objeto quanto adjunto adverbial; as PCs focalizam apenas sujeito; e as CLIV-invs focalizam sujeito.[26]

Podemos apontar como semelhanças entre nosso estudo e o de Braga os seguintes fatos: as CLIVs focalizam as três funções apontadas, embora nossos resultados revelem uma preferência pela focalização de objeto; as PCs, em nossa pesquisa, também focalizam preferencialmente sujeito, mas não exclusivamente.

No estudo de Kato et al. (1996, p.335), com dados do PB culto (Projeto NURC), os resultados revelam que as PCs focalizam predominantemente o objeto direto (63%), e as CLIVs focalizam tanto sujeito (36%) quanto objeto direto (36%); e as CLIV-invs focalizam mais sujeito (50%) e adjunto (41%).[27] As autoras argumentam que as PCs têm características discursivas e sintáticas ideais para focalização de objeto direto: discursivas, por focalizarem apenas elemento novo no discurso, função discursiva prototípica do complemento verbal; e sintáticas, por obedecerem à ordem canônica do PB atual SVO, mantendo o objeto direto após o verbo. Esclarecem que muitos casos encontrados de PCs que são classificados como focalização de sujeito, na verdade, ocorrem com verbos ergativos, demonstrando que o foco é argumento interno do verbo. Explicam que essa característica das PCs de focalizarem predominantemente objeto direto possibilita o apagamento e a não identificação do pronome relativo, estratégia que faz surgir a denominada PC-red.

Ao contrário dos resultados encontrados no estudo de Kato et al. (1996) sobre o PB culto, nos dados de nossa pesquisa a focalização de sujeito favorece o uso de PC, tal qual o resultado encontrado em Braga (1991, p.114). Observamos também que todos os exemplos de PC citados por Braga (em que as PCs só focalizam sujeito) não ocorrem com verbos ergativos, como mostram os exemplos em (46), assim como a maioria dos exemplos de PC documentados em nosso estudo, exemplos em (47):

(46) a. **Quem** fica pedindo pão **é** *mendigo lá na porta*. (BRAGA,1991, p.120)

b. **Quem** que me ensinou **foi** *o cara lá da minha avó*. (BRAGA,1991, p.117)

c. **Quem** segura **é** *o Sandro*. (BRAGA,1991,p.111)

[26] A autora acrescenta a focalização de tópico frasal pelas CLIVs; contudo, por não considerarmos esta opção nesta pesquisa, não a comentamos.

[27] As autoras só apresentaram os números e os percentuais desses três tipos de sentenças clivadas: CLIV, PC e CLIV-inv, embora tenham mostrado as ocorrências de outros tipos.

(47) a. **Quem** compô a fazenda aí **foi** *o finado Ségio*. (CZ-10)

b. **Quem** fez tudo **foi** *ele*. (CZ-12)

c. **Quem** sabe **são** *eles*. (SP-06)

Diante desses resultados, atribuímos as diferenças e semelhanças entre os resultados aqui apresentados e os de Kato et al. (1996) e de Braga (1991) a contextos discursivos que possam ter favorecido ou não uma preferência de focalização de uma determinada função sintática nas PCs, uma vez que essa estratégia se mostrou capaz de focalizar tanto sujeito quanto objeto.

Um fator extralinguístico que se mostrou bastante interessante em nossa análise dos dados foi a variável faixa etária, em relação à CLIV-sem-cóp. Os resultados estão na Figura 2.

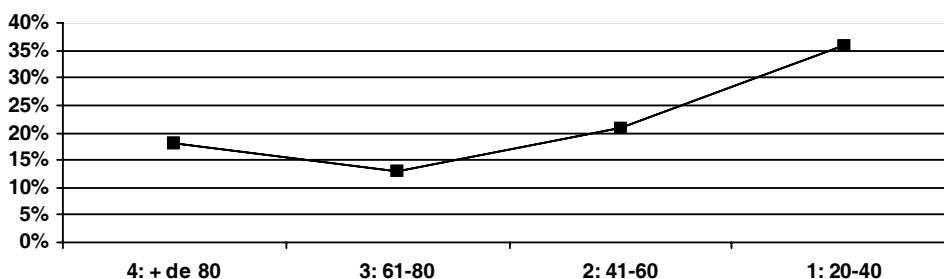



Figura 2 - Atuação da variável faixa etária na estratégia de CLIV-sem-cóp
Fonte: Côrtes Junior, 2006, p.110.

Vemos que a frequência da CLIV-sem-cóp apresenta um percentual de 18% na faixa 4, uma redução de 5 pontos percentuais na faixa 3, e implementa um crescimento de 8 pontos percentuais a mais na faixa 2, chegando a uma frequência de 36% na faixa 1. A ordem crescente de uso da CLIV-sem-cóp, partindo da faixa mais velha, mostra que é uma estratégia antiga na comunidade e que amplia sua realização nas faixas mais novas.

8.7 Para finalizar, um pouco de estrutura e de história

Falta-nos fornecer respostas para as perguntas colocadas na seção 8.1, sobre as diferenças entre as variedades portuguesa e espanhola americanas, de um lado, e a variedade espanhola peninsular, do outro, bem como procurar uma explicação para as diferenças entre PE e PB. A questão é saber quais evidências os dados das comunidades de afrodescendentes fornecem para entender a variação observada nas seções acima.

Do ponto de vista diacrônico, as duas variedades peninsulares realizavam construções clivadas no século XVI, quando então começaram a serem implantadas na América românica:

(48) a. En una escalera *fue que* reñimos (MORENO CABRERA, 1999, p.4281)

b. O voto *é que* se jogue com tôda a baralha (LOBO-1579[28])

Desse modo, espera-se que os dois dialetos possam ter influenciado no desenvolvimento de construções clivadas nos dialetos americanos, pois fazia parte do *input* oferecido pelos colonizadores. Consideramos, assim, que as questões sobre contato só seriam apropriadas caso os dialetos americanos não tivessem desenvolvido construções de clivagem. É relevante acentuar, portanto, que conhecer fatos da história dos dialetos transplantados para a América do Sul é crucial para entender as influências do multilinguismo generalizado na formação histórica do PB.

As diferenças entre PE e espanhol peninsular moderno resultam de desenvolvimentos gramaticais independentes, que levaram à obsolescência desse tipo de construção no espanhol peninsular. Na história das duas línguas, as construções de focalização do tipo que figura em (48) estavam em variação com uma outra estratégia de focalização, a focalização V2, como ilustrado nos exemplos a seguir:

(49) a. MANZANAS compró Pedro (y no peras) (ZUBIZARRETA, 1999, p.4239)

b. SÓ ESSES CDs me recomendou Maria (KATO; RAPOSO, 1996, p.269)

A ordem em que os constituintes ocorrem mostra que o foco (constituinte com a função de objeto) antecede o verbo, e o sujeito o segue; chama-se focalização V2 por o verbo ocupar linearmente a segunda posição (entre outras questões estruturais). As diferenças entre o PE e o espanhol encontram então uma explicação quando se observa que o espanhol peninsular continua a realizar a focalização V2, enquanto o PE moderno só apresenta uso residual dessa construção. Isso significa dizer que as focalizações com *é que* são desenvolvimentos históricos das focalizações V2. Desde o português arcaico que se observa a variação entre as duas formas de focalizar, tendo vencido a disputa a forma inovadora com *é que*, embora a variação tenha durado muitos séculos.

De modo semelhante, oscilações em construções interrogativas, também uma construção de focalização, entre uma interrogativa V2 e uma interrogativa com *é que* é observada na documentação do PE clássico:[29]

---

[28] Data do nascimento do autor. *Corpus* Histórico do Português Tycho Brahe. IEL, Unicamp. Projeto coordenado por Charlotte Galves.

[29] Dados recolhidos na gramática do padre Dom Jeronymo Contador de Argote, nascido em 1676. *Corpus* Histórico do Português Tycho Brahe. IEL, Unicamp.

(50) a. E quantas castas de palavras tem a língua Portuguesa, e as suas Orações?

b. E quando *é que* são relativos?

Também nessa variação, venceu a forma gramaticalizada *é que*, em que a cópula não tem valor verbal e *que* não é um introdutor de sentença subordinada. As interrogativas V2 são, no PE moderno, formas de escrita culta, mesmo assim de uso restrito (KATO; MIOTO, 2005). Os dados de clivadas com dupla cópula (cf. exemplos em (37)-(39)) favorecem esta análise, a de gramaticalização de *é que*.

Assim, a resposta à pergunta (i) considera que as duas formas de construir focalizações já estavam em variação nos dialetos peninsulares que abordaram às terras hoje brasileiras. Nos dialetos românicos americanos e no PE, prevaleceu a forma inovadora, com *é que*; o espanhol peninsular continuou com a focalização V2, para declarativas e interrogativas, tornando as clivadas obsoletas.

A *clivada sem cópula* é a construção inovadora no PB (cf. dados em (33)), presente em qualquer variedade dialetal do PB, mas ausente do PE culto e rural. Consideramos que este tipo de clivada resulta de uma reanálise da construção de recomplementação ilustrada em (34)-(36).[30] Mascarenhas (2007) argumenta que, no PE, o constituinte ensanduichado entre os dois *que* tem uma leitura discursiva única, a de tópico. No PB, ao contrário, a leitura discursiva é ambígua, recuperando mais facilmente o constituinte ensanduichado como foco do que como tópico. Os julgamentos que definimos para os exemplos a seguir:

(51) a. João disse **que** *Ana* **que** foi ao cinema (e não Maria / *e não ao teatro)

b. João disse **que** *Ana* **que** *ela* foi ao cinema (e não ao teatro / *e não Maria)

indicam que a leitura de tópico do constituinte ensanduichado só é possível se houver um pronome retomando o constituinte (cf. exemplo 51b); se não houver retomada pronominal, a única leitura possível é a de foco (cf. exemplo 51a). Assim, na aquisição do português, aconteceu a reanálise do segundo *que* como um marcador de foco, opção não disponível ao PE. A partir do momento da reanálise, a variação entre clivada com e sem cópula passa a fazer parte do dialeto brasileiro. A Figura 2 indica que a reanálise é antiga, pois ocorre na fala dos mais velhos em 18% dos dados.

Além disso, essa é uma reanálise que se difunde pelo PB em geral. Embora não seja específica das comunidades de afrodescendentes, pode ter delas se originado.

[30] As construções de recomplementação deixam de ser atestadas na escrita portuguesa a partir do século XVI; contudo, continuam a ser usadas em escrita informal (cf. exemplos em 36) e na fala rural (cf. exemplos em 34) e mesmo na fala de informantes cultos (MASCARENHAS, 2007).

9 As construções de tópico

*Edivalda Araújo**

Este capítulo tem por objetivo analisar os tipos de construções de tópico encontrados na gramática das comunidades rurais afro-brasileiras isoladas, comparando-os com as construções atestadas por Pontes (1986, 1987) e Galves (1998, 2001) no português brasileiro urbano. O capítulo está estruturado como segue: na seção 9.1, apresentamos uma interpretação teórica do tópico, sob a perspectiva da gramática gerativa, na sintaxe, e da estrutura da informação, no discurso;[1] na seção 9.2, definimos a metodologia utilizada na realização dos inquéritos selecionados para estudo e também os procedimentos para a análise linguística; na seção 9.3, analisamos os tipos de tópico encontrados no *corpus* da pesquisa; ao final, na seção 4, estabelecemos a relação entre os dados linguísticos e os dados sociais, quais sejam: estada fora da comunidade por um período mínimo de seis meses, nível de escolarização e idade; também discutimos a provável influência do contato do português com as línguas africanas em relação às construções de tópico.

9.1 Panorama teórico

(1) *O carro num podia tirá*. (CZ-6)[2]

À primeira vista, frases como esta provocam certo estranhamento, uma vez que somos guiados pelo senso comum de que o verbo *tirar* exige um sujeito com o

* Agradecemos à profa. Dra. Ilza Ribeiro pela análise e indicações feitas ao corpo do trabalho. Os problemas que persistirem são de nossa responsabilidade.

[1] A palavra *discurso* ou *discursivo/a* está sendo usada neste texto na acepção da enunciação, não estando direcionada à perspectiva da Análise de/o Discurso.

[2] Exemplo retirado do *corpus* em análise. A sigla indica a comunidade de Cinzento, e o número indica o informante. Para as demais comunidades analisadas neste capítulo, cf., a seguir, a seção 9.2. As comunidades rurais afro-brasileiras analisadas no Projeto *Vertentes* são descritas no capítulo 2 deste livro, e a metodologia empregada na constituição das amostras de fala vernácula é apresentada no capítulo 5.

traço [+agente], um objeto direto e também um locativo (*alguém* tirar *alguma coisa* de *algum lugar*). Seguindo essa perspectiva, a análise sintática da frase acima evidencia dois fatos: (i) a ausência de um sujeito com o traço [+agente] e do locativo; e (ii) o sintagma nominal que inicia a frase — *o carro* — não pode ser o sujeito, porque não porta o traço semântico [+agente]. Uma leitura desse tipo, entretanto, se centra na estrutura sintática, levando em consideração apenas a seleção de argumentos do verbo.

Um outro olhar para essa frase revela uma análise diferente, se levarmos em consideração não só a estrutura sintática, mas também a estrutura da informação. Em uma realização canônica dessa frase, teríamos a seguinte construção:

(2) X *num podia tirá o carro*.

Observamos em (2) que o sujeito do verbo está indeterminado (representado por X), mas o objeto direto foi realizado — *o carro*. Nesse lugar canônico, à direita do verbo, *o carro* atende aos requerimentos sintáticos; mas, para atender aos requerimentos da estrutura da informação, esse sintagma nominal é deslocado para a posição de tópico, no início da oração, resultando na frase realizada pelo informante: *O carro num podia tirá*. Ou seja, *o carro* é um elemento proeminente, faz parte de um conhecimento partilhado entre ouvinte e falante,[3] funciona, então, como o tópico da oração.

Sob a perspectiva sintática, o tópico é visto como um sintagma nominal, lexical ou pronominal que se realiza numa posição geralmente deslocada à esquerda, na camada externa da oração, o CP (RIZZI, 1997), em torno do qual é construído um predicado ou comentário. Mas, precisamente, o fato de ele se realizar na camada do CP implica que ele tem propriedades discursivas, porque é aí que ocorre a interface sintaxe-discurso e, consequentemente, a verificação dos traços discursivos (RIZZI, 1997, 2003; CHOMSKY, 2001a). No que se refere à localização sintática do tópico, assumimos, juntamente com RIZZI (1997, 1999), que a oração dispõe de três camadas: a lexical, onde os elementos sofrem a primeira inserção na computação sintática, o *v*P; a flexional, para onde os elementos são movidos para se verificarem os seus traços morfológicos, o IP (ou TP); e a discursiva, o CP, para onde os elementos são movidos para se verificarem os traços discursivos, como o de tópico, o de foco, o de escopo de interrogação, entre outros.

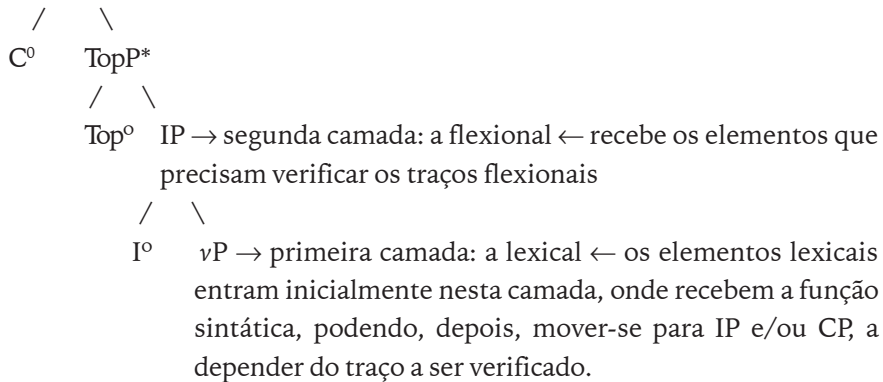
Dentro dessa perspectiva, a estrutura de análise que adotamos para os elementos na posição de tópico, de forma resumida, é a seguinte:

[3] Considerando-se que se trata de um *corpus* oral, faremos referência apenas ao par falante/ouvinte; mas as construções de tópico não se restringem apenas à realização oral — podem também aparecer em textos escritos.

(3) CP → terceira camada: a discursiva ← recebe os elementos que precisam verificar os traços discursivos

C⁰ TopP*

Top° IP → segunda camada: a flexional ← recebe os elementos que precisam verificar os traços flexionais

I° *v*P → primeira camada: a lexical ← os elementos lexicais entram inicialmente nesta camada, onde recebem a função sintática, podendo, depois, mover-se para IP e/ou CP, a depender do traço a ser verificado.

Desse modo, a menção à posição do tópico implica considerá-lo em TopP, podendo ou não ser retomado por um elemento em posição mais baixa que a dele, em IP ou *v*P, como veremos na análise dos tipos de tópico na seção 9.3.

Sob a perspectiva da estrutura da informação, o tópico é visto como um direcionamento do discurso, sinalizando que o falante pressupõe ser esse constituinte uma informação conhecida pelo ouvinte. Porque vem no início da oração, o tópico tem a função de orientar o ouvinte para a construção do significado ou para o estabelecimento de relações com outras informações na sentença, no texto ou na situação. A identificação de um tópico, de acordo com essa perspectiva, depende da sua localização em um determinado contexto, uma vez que a sua função é indicar sobre o que se está falando. Portanto, a noção de tópico só pode ser propriamente definida em termos das relações entre uma sentença e o contexto em que ela ocorre. A abrangência do tópico, desse modo, envolve tanto a frase quanto o discurso, uma vez que, embora seja deslocado à esquerda, o tópico mantém com a oração seu papel ativo na estrutura sintático-semântica da mesma.

9.2 Metodologia

Para constituir o *corpus* aqui analisado, foram selecionados 26 inquéritos com afro-brasileiros de quatro comunidades rurais isoladas do interior do Estado da Bahia, com a seguinte distribuição: 07 inquéritos de Helvécia (HV); 07 de Cinzento (CZ); 06 de Rio de Contas (RC); e 06 de Sapé (SP).[4] As variáveis sociais controladas na análise foram: idade, nível de escolaridade e estada fora da comunidade.

[4] Uma descrição dessas comunidades é apresentada no capítulo 2 deste livro, e a metodologia empregada na constituição das amostras de fala, no capítulo 5.

A análise das construções de tópico no *corpus* em questão parte da caracterização do tópico como um sintagma nominal definido, identificável, ativo e referencial, realizado por um nome ou pronome,[5] sendo uma informação partilhada pelos participantes da conversa. Seguindo essa orientação, os tópicos detectados se realizaram ou pela retomada da fala do documentador, como no exemplo em (4), ou pela retomada de um elemento já citado anteriormente pelo informante, como no exemplo em (5):

(4) *Seu pa, sua mãe... também nasceram em Cinzento?*

É... minha mãe eu num... num lembro bem de onde ela nasceu não. (CZ-3)

(5) *É... o bicho... hum... hum... É, sim... o caçadô ficava perdido, mas num era... num era o ININT era uma... era uma folha... [...] É. Aquela folha... os menino saía, ia caçá, né, aí bateu aqui nos óio do cachorro... o cachorro ele é meio...* (HV-7)

Em (4), vemos na fala do informante a retomada do elemento linguístico da fala do documentador (*minha mãe / sua mãe*). Em (5), o informante cita *folha* e faz retomada posterior. Em ambos os casos, os elementos retomados podem ser considerados tópicos, porque já fazem parte do conhecimento partilhado entre os falantes.

A análise do tópico no *corpus* seguiu, então, esses procedimentos: identificação do elemento como retomada do que foi dito antes, ou sinalização de que o elemento inicial estava sendo usado como orientador discursivo, sobre o qual seria feito um comentário, eliminando-se os casos em que pudesse haver ambiguidade de análise com o foco, por exemplo.

9.3 Análise dos dados linguísticos

Em termos gerais, foram computadas 267 realizações de construções de tópico, distribuídas em oito tipos diferenciados, como pode ser visto na Tabela 1.

Tabela 1 - Quantificação dos tipos de tópico

TIPOS DE TÓPICO	QUANTIDADE	%
Topicalização de objeto direto	61	23
Tópico pendente com retomada	52	19
Tópico cópia	50	19
Tópico sujeito	29	11
Tópico pendente	28	10
Duplo sujeito	23	9
Topicalização selvagem	19	7
Locativo	5	2
Total	267	100

[5] Regido ou não por uma preposição.

A realização diferenciada desses tipos de tópico será explicada a seguir, na ordem dos percentuais de realização mostrados na Tabela 1.

9.3.1 Topicalização de Objeto Direto (TOD)

Um dos tipos de tópico mais realizados é a Topicalização de Objeto Direto,[6] com 23% do total de ocorrências. Esse tipo de construção, como se pode ver na amostra de casos de TOD presentes abaixo transcrita, apresenta as seguintes características: (i) um objeto direto deslocado à esquerda sem retomada clítica interna à oração (6-11); (ii) o sintagma nominal que compõe este objeto normalmente é definido (quer dizer, acompanhado por um determinante definido, em (6-7)); (iii) não sofre restrições de ilhas (8-9); e (iv) pode ocorrer em contextos de encaixada (10-11):

(6) *os pé de café trocô.*[7] (HV-20)

(7) *A cachaça eu bebo todo dia, se eu todo dia eu fô lá na praça.* (RC-8)

(8) *... esses criatório (porco, galinha) tamém eu tem muitos ano qu'eu num crio, né?* (HV-12)

(9) *certas coisa tem hora que num pode vendê mais caro, o preço é aquele, né?* (RC-26)

(10) *É mais perigoso, porque uma... uma mulé... grávida, pá ganhá neném em casa, eu acho muito perigo, porque lá no hospital... já é ôta coisa, né?* (HV-4)

(11) *Parece que o poquim que ocê aprende na escola que dorme, quande é no ôto dia parece que já num sabe mais, puquê tanto pá fazê.* (CZ-1)

No que se refere ao *corpus* do português rural afro-brasileiro, é importante ressaltar que, no cômputo geral, essas construções sem retomada clítica são as mais usadas (cf. Tabela 1). É preciso salientar que a TOD não é construção exclusiva do português rural afro-brasileiro, mas está presente em todas as modalidades do português brasileiro. Inclusive, podemos considerar, a partir do texto de Cyrino (1993), que a TOD tem registro no português brasileiro escrito desde o século XIX, conectada à mudança em direção à perda do clítico acusativo de terceira pessoa, sendo a sua relação com a posição interna na oração licenciada no contexto discursivo (cf. seção 9.4).

---

[6] Alguns autores estabelecem relação entre esse tipo de construção e as de objeto nulo, uma vez que a posição interna deixada por esse elemento deslocado para a posição de tópico é foneticamente nula: não há um pronome visível para retomar o objeto topicalizado. Para mais informações sobre o objeto nulo, ver o capítulo 17 deste livro.

[7] É preciso ressaltar que no exemplo em (6) o informante não flexionou o verbo na primeira pessoa do singular, prática comum nos falantes mais velhos dessa comunidade. Em outros informantes, a forma realizada seria: *troquei*.

9.3.2 Tópico Pendente com Retomada

Dentre as realizações das construções de tópico, o Tópico Pendente com Retomada obteve um percentual de 19%. Este tipo de tópico mantém uma relação semântica com a oração porque é retomado por um elemento interno à oração, como um pronome forte ou clítico, uma expressão genérica, uma categoria vazia, um pronome demonstrativo, um numeral, dentre outros.

Nos casos analisados, ocorreu a retomada por: pronome tônico, quantificador, numeral, pronome demonstrativo e categoria vazia, como se pode ver na distribuição da Tabela 2.

Tabela 2 - Tipos de retomada nas construções de Tópico Pendente com Retomada

CONSTRUÇÕES DE TÓPICO PENDENTE COM RETOMADA	QUANTIDADE	%
Relação continente / contido	10	19
Pronome correferencial	8	15
Retomada vazia na posição de sujeito	7	13
Pronome demonstrativo	6	12
Quantificador	6	12
Numeral	5	10
Pronome pessoal na posição de objeto	4	8
Retomada pronominal na posição de sujeito - oração subordinada	4	8
Pronome indefinido	2	4
Total	52	100

De acordo com os dados numéricos presentes na Tabela 2, observamos que o tipo de retomada mais frequente é o da relação semântica continente/contido, com 19%. Nesse tipo de retomada, existe geralmente uma relação do mais amplo para o mais específico, ou seja, o elemento citado no tópico é um termo semanticamente mais amplo, contendo o termo que o retoma internamente à oração, como é evidenciado a seguir:

(12) *Mas, meus porco, você pricisa de vê, quand'eu crio um... **um leitão*** (HV-20)

(13) *Jogo, naquele tempo **o futebol** era mais efetivado ainda.* (CZ-3)

À retomada continente/contido segue-se a retomada por pronome correferencial, ou seja, pronomes que retomam pronomes com o mesmo índice referencial. Esse tipo de retomada, exemplificada logo abaixo, representa 15% dos casos (cf. Tabela 2).

(14) *eu, dependeno de **mim**, e dependeno dessa qu'eu tenho, ela disse que: "óia Vardo por **mim**... tua ex-mullhé pode chegá aqui pa conversá **comigo**..."* (SP-12)

(15) *Eu, **meu** nascimento é daqui mesmo, **minha** residênça é aqui*. (CZ-9)

A seguir, encontramos a retomada do tópico por um elemento foneticamente nulo na posição de sujeito, representando 13% dos dados. Essa retomada se caracteriza pelo deslocamento de um elemento da posição de sujeito tanto de oração principal (em (16)) quanto de subordinada (em (17)), ficando em seu lugar um elemento nulo. Os seguintes exemplos ilustram isso:

(16) *Aquela folha*... os *menino saía, ia caçá, né, aí* **Ø** *bateu aqui nos óio do cachorro* (HV-7)

(17) *A cabra diz que* **Ø** *é criada com cerca bem feitcha*... (CZ-12)

Encontramos também a retomada por pronome demonstrativo, em 12% dos dados, podendo haver repetição anafórica do elemento topicalizado seguindo o pronome, como em (19):

(18) *Mas esses tratô que a gente tem aqui*… ***esses*** *daqui… é… quato macha, cinco* (HV-4)

(19) *a oração que a gente orô pra ela*, ***essa oração****, eu... eu digo mermo que serviu mermo ela*. (SP-5)

Os outros tipos de retomada, como a de quantificador (20), numeral (21), pronome pessoal na posição de objeto (22), retomada pronominal na posição de sujeito de encaixada (23) e pronome indefinido (24), tiveram uma realização mais baixa em comparação com as discutidas acima (cf. Tabela 2):

(20) *Parto que tivé agora*, ***todos*** *vai... vai... tem que... atendeu a precisão já manda atrás de um carro logo, ó*. (CZ-9)

(21) *Irmãos, ieu teve* ***trêi***. (HV-13)

(22) *A cana... cê prantô* ***ela****... ela brotô*. (RC-8)

(23) *Minha mãe… acho qu'****ela*** *tá com mais de trinta e tanto ano*! (HV-13)

(24) *Aí a gente*... ***um*** *vai e os* ***outro*** *fica isperando*. (CZ-3)

O mais interessante na construção de Tópico Pendente com Retomada é que a retomada do tópico pode ocorrer em qualquer posição interna à frase. Além disso, existe uma relação semântica entre os elementos e também uma "fraca" relação sintática, porque não há correspondência morfológica obrigatória entre o elemento topicalizado e o termo que o retoma internamente à oração; isto é, é possível que o elemento topicalizado esteja no plural, mas a sua retomada esteja no singular (exemplo em 12), ou no singular e a sua retomada no plural, como em (20).

9.3.3 Tópico Cópia

As construções de Tópico Cópia, com 19% das realizações (cf. Tabela 1), recebem essa denominação porque, diferentemente do tipo analisado acima, a retomada interna à oração é feita pela cópia do termo topicalizado, como se pode ver em (25-26):

(25) *aí o tratô… a carreta empurrô **o tratô**, e aí desceu de ladêra abaxo lixado…* (HV-4)

(26) *agora, Teofil'Otone, num conheço **Teofil'Otone** direito...* (HV-12)

Evidência de que esses elementos estão na posição de tópico é o fato de, em alguns casos, a sua retomada ocorrer por um elemento focalizado[8] (foco informacional ou contrastivo), como ocorre em (26).

9.3.4 Tópico Sujeito

Essa construção, com realização de 11% (cf. Tabela 1), caracteriza-se por: (i) apresentar um sintagma preposicional, locativo ou adjunto, deslocado à esquerda sem a preposição, comportando-se como o sujeito da frase porque estabelece concordância com o verbo; (ii) não haver um pronome lembrete retomando o sintagma nominal anteposto; e (iii) não existir concordância verbal entre o verbo e o seu argumento externo, que se realiza em posição pós-verbal, conforme demonstram Pontes (1987) e Galves (1998, 2001).

Nessas construções, o sintagma nominal move-se de posições internas da oração, mas parece não conseguir alcançar a posição de TopP, na camada do CP.[9] Nesse caso, como em (27a), o sintagma nominal parece ficar interno ao IP, o que é evidenciado pelo fato de ele poder desencadear concordância com o verbo,[10] como se fosse o sujeito da oração. Além disso, o argumento selecionado pelo verbo está posposto (*dentes*), mas não concorda com o verbo porque a frase resultaria agramatical, como se vê abaixo, em (27b):

(27) a. A *Sarinha* está nascendo dentes. (PONTES, 1987, p.35 (90)).

b. *A Sarinha estão nascendo os dentes.

A frase em (27a) pode ser ambígua em relação à concordância do verbo (cf. nota 10), mas não a frase em (28), ouvida em um programa de rádio:

(28) a. Os *jogadores* estão crescendo o cabelo.

b. *Os jogadores está crescendo o cabelo.

As frases acima, em (27b) e (28b), como se pode ver, ficam agramaticais com a concordância do verbo com o sujeito posposto, em função de os sintagmas

---

[8] A respeito das construções de foco, veja-se o capítulo 8 deste livro.

[9] Como foi representado na estrutura em (3).

[10] É preciso considerar a ambiguidade da oração: pode ser que o falante não tenha realizado nenhum tipo de concordância. Ou seja, em função do enfraquecimento da flexão verbal do português brasileiro, é possível que o falante tenha realizado uma construção em que o sintagma nominal esteja na posição de tópico, mas o verbo esteja sem concordância de pessoa (GALVES, 1998, 2001). Ex: $[_{CP}$ $[_{TopP}$A Sarinha $[_{IP}$ $\varnothing$ está nascendo dentes.]]]

nominais iniciais, *a Sarinha* e *os jogadores*, ficarem "soltos", como se não fizessem parte da construção da frase. Essas construções são consideradas de Tópico Sujeito, justamente porque é um tópico, e não o sujeito, que estabelece concordância com o verbo. No português brasileiro urbano, como apontado por Pontes (1986), essas construções estão presentes inclusive na fala culta. Contudo, registros desse tipo não foram encontrados em outras línguas românicas (GALVES, 1998), o que torna o português brasileiro *sui generis* em relação a essas construções de tópico.

No caso do português afro-brasileiro, também foram encontradas construções do Tópico Sujeito, que distribuímos em três grupos:

Grupo I: o verbo é inacusativo,[11] o que implica que ele seleciona um argumento interno, mas esse argumento não desenvolve concordância com o verbo (como foi visto nos exemplos em (27b) e (28b) acima). Em ambos os casos, a primeira leitura da frase leva a analisar o sintagma nominal inicial como se fosse o sujeito da oração. É o que encontramos nos exemplos a seguir, em que o sintagma nominal inicial provém de outras posições internas à oração: adjunto adnominal (29), complemento nominal (30), adjunto adverbial (31) e objeto direto do verbo existencial (32):

(29) *o carro afundô as roda...* (CZ-6) [as rodas do carro afundaram]

(30) *a desculpa do Nordeste tem muito interesse.* (CZ-6) [há muito interesse na desculpa do Nordeste]

(31) *Ago'lacraia é trabáio perdido.* (RC-8) [é trabalho perdido com lacraia]

(32) *o cacau tinha mais que o guaraná,* (SP-5) [tinha mais o cacau que o guaraná]

Admitimos que os exemplos acima são ambíguos em relação à concordância do verbo com o tópico, em função de ambos estarem no singular, com exceção do exemplo em (29). Neste, é possível identificar que o provável sujeito do verbo, *as roda*, está no plural, mas o verbo manteve-se no singular, em concordância com o tópico.

Grupo II: os verbos selecionam dois argumentos, um interno e um externo, mas o interno é alçado para a posição de tópico e, provavelmente, desencadeia concordância com o verbo, enquanto a posição do argumento externo está vazia,

[11] Os verbos inacusativos são aqueles que selecionam apenas argumentos internos, podendo estes serem alçados para a posição de sujeito, estabelecendo a concordância com o verbo. Geralmente, esses verbos têm os chamados sujeitos pacientes, como o verbo *nascer*, *morrer*, *crescer*, os existenciais e os de ligação, entre outros. Uma frase com verbo inacusativo seria assim gerada: *nascer João*, porque este verbo seleciona um argumento interno, *João*. Esse sintagma nominal depois seria movido para a posição de sujeito para estabelecer a concordância com o verbo, resultando em: *João nasceu*. Com os verbos existenciais, em geral, não ocorre o alçamento do seu argumento interno para a posição de sujeito, porque esta fica preenchida com um expletivo nulo.

realizada por um sujeito nulo não referencial.[12] É o que pode ser observado nos exemplos abaixo, em que o objeto direto (33), o adjunto adverbial (34) e o sujeito do infinitivo (35) ocupam a posição de Tópico Sujeito:

(33) *essa casa foi ligêro... que o moço... sabia trabaiá um pôco...* (CZ-12) [foi ligeiro (fazer) essa casa]

(34) *Quaqué terra faz adôbo adôbo.* (CZ-11) [faz-se adobe com qualquer terra]

(35) *ele já foi preciso ficá ficá internado...* (HV-1) [já foi preciso ele ficar internado]

Grupo III: não há nenhuma relação do tópico com a posição interna na oração. Ou seja, o elemento que aparece como provável sujeito do verbo não faz parte da sua grade temática e não tem relação semântica com ele. Consideramos, nesses casos, que houve a inserção direta de um tópico na posição de Spec, SujP,[13] em IP, sem que o sintagma tivesse sofrido algum tipo de deslocamento de algum lugar interno da oração, daí analisá-lo diretamente como tópico. É o que pode ser visto nas frases abaixo:

(36) *a residência deles é nascido, criado tudo aqui mesmo.* (CZ-9)

(37) *o canivete é coisa que uma unha tava suja... [melada] aqui... eu fui rapá o canto da unha!* (CZ-11)

Algumas dessas construções que denominamos Tópico Sujeito são apresentadas por Pontes (1987) em sua análise do português urbano, como as identificadas no Grupo I. Os casos dos Grupos II e III, embora não sejam citados pela autora, foram encontrados nas comunidades em estudo, mas acreditamos estarem presentes no vernáculo brasileiro, em qualquer dialeto.

9.3.5 Tópico Pendente

Nas construções de Tópico Pendente, que obtiveram 10% das realizações (cf. Tabela 1), o sintagma nominal que inicia a oração tem uma relação semântica com a frase, mas não uma relação sintática. Ou seja, não há um lugar interno na oração em que ele possa ser inserido. Em geral, essa construção pode ser introduzida por algumas marcas formais, como *quanto a..., no que se refere a...*, ou por um sintagma nominal sem determinante que o encabece. No *corpus* sob análise, só foram encontradas construções de Tópico Pendente com sintagmas nominais

[12] Isso está em consonância com os estudos realizados por Kato (2002b) e Galves (1998), para quem, no português brasileiro, quanto menos referencial for o sujeito pronominal, menos provável a sua realização.

[13] Explicações sobre essa posição de sujeito estão em Araújo (2006).

iniciais, acompanhados ou não de determinantes (cf. exemplos em (38-39) e em (40-41), respectivamente):

(38) *...no caso, o rapaz aqui..., todo mundo que cria bicho aqui me pede uma força dia de domingo, que nem hoje mesmo, de manhã cedo estava montano animal aí.* (SP-6)

(39) *A farinha... prantô a mandioca, ['maduceu]... leva um ano... um ano e pôco... rancô, chegô na casa de farinha, raspô... relô... botô na prensa... arrochô e tirô a água toda,* (SP-12)

(40) *médico sempre aí nas Serra, nesse Rapa mermo tem um posto...* (SP-4)

(41) *Batuque aí, botava era um... era um... era um cabuêro com tambô, subia em cima, o ôto ia tocá e... as muieres fazia, sambano* (HV-12)

9.3.6 Tópico com Cópia Pronominal ou Duplo Sujeito[14]

Essa construção se caracteriza pelo deslocamento de um sintagma nominal para a posição de tópico e uma retomada pronominal deste sintagma na posição de sujeito interna à oração, conforme exemplos abaixo:

(42) *A sussuarana,* ***ela*** *pensa carnêro tá no mato, que... que 'cê num tocô, elas vai no rebanho e mata* (CZ-12)

(43) *E aqueles mele,* ***eles*** *é que faz arco.* (RC-26)

A presença obrigatória do pronome na posição do sujeito nessas construções, de acordo com Galves (1998), justifica-se em função de o verbo ter perdido o traço de [pessoa] no português brasileiro, daí a necessidade de marcação desse traço no sujeito.[15] Construções como essas, apesar de abundantes no português coloquial urbano, foram encontradas no *corpus* do português rural afro-brasileiro com um baixo percentual (9% do cômputo geral, de acordo com a Tabela 1).

9.3.7 Topicalização Selvagem

A Topicalização Selvagem, de acordo com Brito, Duarte e Matos (2003),[16] ocorre quando há deslocamento de um PP, objeto indireto, desacompanhado de

---

[14] Cf. designação de Duarte, 1993.

[15] Não vou entrar na discussão da posição do sujeito, se interno ou externo à oração. Para discussões a esse respeito, cf.: Costa (2000, 2001); Galves (2001); Costa e Galves (2000); Barbosa, Kato e Duarte (2001).

[16] O *corpus* de análise dessas autoras é o português europeu urbano.

preposição.[17] Diferentemente dessas das autoras, consideramos que outras funções sintáticas (regidas por preposição, mas realizadas sem ela) possam aparecer nesse tipo de construção. Nesse sentido, entram nessa classificação, além dos objetos indiretos (44), os seguintes casos: complementos nominais (45), agente da passiva (46), adjunto adverbial (47), conforme os exemplos encontrados no *corpus*:

(44) *Futebol, a gente brincava, né...* (HV-4)

(45) *As região vizinha eu sô o presidente e tenho o direito de pedi, revindicá os pedido.* (CZ-6)

(46) *Ah, lacraia já fui mordido muitas vez.* (RC-8)

(47) *São Paulo, eu trabaiei... eu... eu trabaiei numa usina... lá pra fora [do] interiô, mas foi cá no estado de Som Paulo mesmo.* (RC-26)

É possível verificar, nos exemplos acima, que os elementos topicalizados são provenientes de lugares diferenciados internos à oração, mas sem a regência da preposição. Essas construções tiveram um percentual de apenas 7% no cômputo geral.

9.3.8 Tópico Locativo

O elemento aqui topicalizado é um locativo, que funciona como adjunto ou de verbos existenciais (48) ou de verbos tradicionalmente considerados intransitivos (49-50). Nessas construções, em geral, a preposição acompanha o sintagma nominal deslocado:

(48) *nesse sertão nosso aqui tem cascavel demais.* (CZ-6)

(49) *pra Conquista ela sempe vai, mais eu... mas só.* (CZ-1)

(50) *No eucalipto, eu trabalhei... no tempo da PLANIBA.* (HV-12)

Essas construções, que obtiveram um percentual muito baixo de realização, apenas 2% (cf. Tabela 1), diferenciam-se da Topicalização Selvagem por apresentarem o sintagma nominal acompanhado de preposição. Supomos que a baixa realização se deva justamente à presença da preposição, uma vez que, de acordo com os dados citados acima em relação ao Tópico Sujeito e à Topicalização Selvagem, parece haver uma tendência, na oralidade, para o apagamento da preposição.

[17] Essa preposição deve ser funcional, sem conteúdo semântico. Podemos ver a diferença nas preposições nos seguintes exemplos: *Gosto de doce* e *Vou com você*. A preposição de, no primeiro exemplo, é considerada funcional, sem conteúdo semântico, porque está ligando o verbo ao seu complemento; o que não acontece com a preposição com, no segundo exemplo. Esta preposição tem conteúdo semântico, porque está inserindo um adjunto adverbial, e não um complemento ao verbo.

9.3.9 Conclusões parciais - dados linguísticos

A análise linguística das construções de tópico revelou mais pontos de confluência com o português brasileiro urbano do que pontos de divergência, sendo esta última relacionada a um tipo de tópico encontrado no *corpus* das comunidades em estudo, o Tópico Cópia. A identificação desse tipo de tópico nessas comunidades não implica ausência no português brasileiro urbano, mas que, provavelmente, ainda não foi citado pela literatura que trata desse assunto. O que queremos ressaltar é que, a partir das comparações realizadas, não há diferenças consubstanciais entre as realizações de tópico do português rural afro-brasileiro e as do português urbano.

Voltando aos dados da Tabela 1, vemos que as construções de tópico mais realizadas no *corpus* foram a de Topicalização de Objeto Direto (23%), a de Tópico Pendente com Retomada (19%) e a de Tópico Cópia (19%). O mais interessante é que essas três realizações revelam movimentos contrários. Quer dizer, a primeira, diferentemente das outras duas, é resultado de uma tendência do português brasileiro em direção a uma retomada vazia de tópico na oração, em função da eliminação, na fala, dos clíticos acusativos de terceira pessoa. As outras duas construções, entretanto, evidenciam uma retomada interna na oração. No caso do Tópico Pendente com Retomada, essa retomada pode ocorrer com qualquer elemento, mas, no caso do Tópico Cópia, a retomada só ocorre com a cópia do elemento topicalizado.

Em seguida, foram encontradas, em número menor, as realizações de Tópico Sujeito e Tópico Pendente, cada uma representando 11% e 10%, respectivamente, dos dados, e as realizações de Duplo Sujeito (ou Tópico com Cópia Pronominal) e Topicalização Selvagem, com representação respectiva de 9% e 7% dos dados. Por último, com pouquíssima realização, está o Tópico Locativo, representando apenas 2% dos dados.

Conforme dados apresentados por outras pesquisas (GALVES, 1998; KATO, 1989), as construções de Tópico Sujeito e de Topicalização Selvagem (esta última nos moldes aqui definidos) caracterizam o português brasileiro, visto que ambas não são encontradas em outras línguas românicas, o que parece indicar interferências linguísticas de outros tipos que podem ter provocado a criação dessas construções nessa variedade do português. A relação dos dados sociais com os dados linguísticos pode revelar alguma indicação dessa interferência, nomeadamente, do contato entre línguas, como será visto na próxima seção.

9.4 Análise da relação dos dados linguísticos com os dados sociais

Nesta seção, procuramos estabelecer uma relação entre os dados linguísticos e os dados sociais, tomando para análise: a permanência ou não na comunidade; o nível de escolaridade; e a faixa etária.

9.4.1 Estada fora da comunidade

Considerando a pesquisa realizada por Pontes (1987) sobre o português urbano, seria de esperar que fosse encontrada uma relação entre estada fora da comunidade e maior número de realização de construções com tópico. A correlação numérica de realização de construções de tópico de pessoas que saíram da comunidade, morando, às vezes, sete anos fora, em outros estados, como São Paulo e Rio de Janeiro, com as de outras pessoas que permaneceram na comunidade, indicou, entretanto, que não há diferença significativa entre a estada fora da comunidade ou não. É o que revelam os dados da Tabela 3.

Tabela 3 - Realizações de tópico relacionadas com a permanência na comunidade ou deslocamento da comunidade

REALIZAÇÕES DE TÓPICO	PERMANÊNCIA NA COMUNIDADE		DESLOCAMENTO DA COMUNIDADE	
	Quantidade	%	Quantidade	%
Topicalização de Objeto Direto	28	**10**	33	**12**
Tópico Pendente com Retomada	30	**11**	22	**8**
Tópico Cópia	26	**10**	24	**9**
Tópico Sujeito	14	5	15	**6**
Tópico Pendente	16	6	12	**4**
Duplo Sujeito	15	6	8	**3**
Topicalização Selvagem	10	**4**	9	**3**
Locativo	2	1	3	**1**
Total	141	**53**	126	**47**

De acordo com os dados da Tabela 3, os falantes que permaneceram na comunidade produziram mais tópico do que os que saíram, com respectivos percentuais de 53% e 47%. Mas, analisando separadamente, temos que as construções de Topicalização de Objeto Direto parecem ter sofrido influência externa, uma vez que há uma diferença percentual mínima de 2% entre a realização dos que se deslocaram da comunidade (12%) e dos que nela permaneceram (10%).

Os tipos Tópico Pendente com Retomada, Tópico Cópia e Tópico Pendente seguem a tendência geral de serem mais produzidos por quem permaneceu na

comunidade. Os tipos Tópico Sujeito, Duplo Sujeito e Tópico Locativo não apresentam diferenças significativas entre permanência e estada fora da comunidade.

Esses dados revelam que a realização das construções de tópico no português afro-brasileiro não sofreu, supostamente, interferência urbana, visto que as pessoas que permaneceram na comunidade foram as que mais produziram construções de tópico. Esses dados, entretanto, não podem ser conclusivos, em função de outras interferências externas, como rádio e televisão.

9.4.2 Nível de escolaridade

No que se refere à realização de tópico e à correlação com o nível de escolaridade, foram encontrados os seguintes resultados presentes na Tabela 4.

Tabela 4 - Tipos de tópico e nível de escolaridade

TIPOS DE TÓPICO	SEMIALFABETIZADO		ANALFABETO	
	Quantidade	%	Quantidade	%
Topicalização de Objeto Direto	29	**11**	32	**12**
Tópico Pendente com Retomada	21	**8**	31	**12**
Tópico Cópia	27	**10**	23	**9**
Tópico Sujeito	16	**6**	13	**5**
Tópico Pendente	15	**6**	13	**5**
Duplo Sujeito	12	**4**	11	**4**
Topicalização Selvagem	9	**3**	10	**4**
Locativo	3	**1**	2	**1**
Total	132	**49**	135	**51**

Os dados numéricos apresentados na Tabela 4 revelam que as pessoas semialfabetizadas produzem menos tópico do que as analfabetas. Separadamente, entretanto, vemos que há algumas construções de tópico que são mais produzidas por pessoas semialfabetizadas, enquanto outras são mais produzidas por pessoas analfabetas. Entre as realizadas em maior número por pessoas semialfabetizadas, encontramos o Tópico Cópia, o Tópico Sujeito e o Tópico Pendente, ao passo que as construções de Topicalização de Objeto Direto foram mais realizadas por pessoas analfabetas. As construções de Duplo Sujeito, Topicalização Selvagem e Locativo não apresentaram diferenças significativas.

O mais interessante é que, ao fazer o cruzamento de dados entre deslocamento ou não da comunidade e escolaridade, encontramos um movimento contraditório, conforme Tabela 5.

Tabela 5 - Relação entre permanência na comunidade ou deslocamento e nível de escolaridade

TIPOS DE TÓPICO	SEMIALFABETIZADOS				ANALFABETOS			
	PERMANÊNCIA		DESLOCAMENTO		PERMANÊNCIA		DESLOCAMENTO	
	Quantidade	%	Quantidade	%	Quantidade	%	Quantidade	%
Topicalização de Objeto Direto	6	2	23	9	22	8	10	4
Tópico Pendente com Retomada	7	3	14	5	23	9	8	3
Tópico Cópia	8	3	19	7	18	7	5	2
Tópico Sujeito	6	2	10	4	8	3	5	2
Tópico Pendente	6	2	9	3	10	4	3	1
Duplo Sujeito	6	2	6	2	9	3	2	1
Topicalização Selvagem	2	1	7	3	8	3	2	1
Locativo	0	0	3	1	2	1	0	0
Total	41	**15**	91	**34**	100	**37**	35	**13**

As pessoas semialfabetizadas que saíram da comunidade, segundo os dados da Tabela 5, produzem mais tópico do que as que permaneceram; enquanto, dentre as pessoas analfabetas, há maior número de realização de tópicos por quem permaneceu na comunidade do que por quem saiu. É um movimento contrário, principalmente se observarmos as construções de Topicalização de Objeto Direto, de Tópico Pendente com Retomada, de Tópico Cópia e de Tópico Pendente. Esses dados, contraditórios, parecem indicar que não se pode atribuir as realizações de tópico nem à interferência externa à comunidade nem à escolarização, inclusive porque, como se pode observar na Tabela 5, os números de realizações entre os semialfabetizados que se deslocaram da comunidade e os de analfabetos que permaneceram são muito próximos, excetuando-se as realizações de Tópico Pendente com Retomada, em que o número de realizações é maior para os analfabetos que permaneceram na comunidade.

O fato de não haver interferência do fator nível de escolaridade nas construções de tópico pode evidenciar que: a) não há muita diferença entre ser analfabeto e semialfabetizado nas comunidades em estudo, o que implica que tanto o analfabeto quanto o semialfabetizado vão apresentar as mesmas tendências para a realização de construções de tópico; b) mesmo que o fator nível de escolaridade tenha alguma influência na fala das pessoas das comunidades estudadas, as construções de tópico, de modo geral, não são alvos de atenção nas aulas de língua (como a concordância, por exemplo), daí não se poder avaliar com precisão se há ou não interferência da escolaridade.

Diante desses resultados, foi necessário inserir mais um dado social, a faixa etária, para se averiguar se realmente há alguma interferência dos dados sociais na produção linguística dos tópicos.

9.4.3 Faixa etária

Analisando a influência da faixa etária na realização dos tipos de tópico, foram encontrados os resultados constantes na Tabela 6.

Tabela 6 - Realizações de tópico por faixa etária

TIPOS DE TÓPICO	FAIXA 1		FAIXA 2		FAIXA 3	
	Quantidade	%	Quantidade	%	Quantidade	%
Topicalização de Objeto Direto	14	**5**	24	**9**	23	**9**
Tópico Pendente com Retomada	10	**4**	16	**6**	26	**10**
Tópico Cópia	15	**6**	15	**6**	20	**7**
Tópico Sujeito	5	**2**	14	**5**	10	**4**
Tópico Pendente	5	**2**	14	**5**	9	**3**
Duplo Sujeito	7	**3**	3	**1**	13	**5**
Topicalização Selvagem	4	**1**	5	**2**	10	**4**
Locativo	1	**0**	2	**1**	2	**1**
Total	61	**23**	93	**35**	113	**42**

Esses resultados evidenciam que, em termos gerais, a faixa 1 realiza menos tópico do que as outras faixas. As faixas 2 e 3 apresentam diferenças de realização de tópico: a faixa 2 realiza mais o Tópico Sujeito e o Tópico Pendente, enquanto a faixa 3 realiza mais o Tópico Pendente com Retomada, o Tópico Cópia, o Duplo Sujeito e a Topicalização Selvagem. Mas as faixas 1, 2 e 3 não apresentam diferenças em relação à Topicalização do Objeto Direto e ao Tópico Locativo.

Excetuando-se as construções de Tópico Sujeito e de Tópico Pendente, que apresentam número maior de construções de tópico na faixa 2, observa-se, em geral, com as outras construções, uma tendência a aumentar o número em direção à faixa 3. Ou seja: a faixa 3 parece realizar mais tópico do que a faixa 2, que, por sua vez, também realiza mais tópico do que a faixa 1. A faixa etária, diferentemente dos outros fatores analisados até então, parece ter influência nas realizações de tópico.

9.4.4 Conclusões parciais - dados sociais

A partir da análise da relação dos dados sociais com os dados linguísticos, pode-se dizer que: (i) não há influência externa para a realização dos tópicos; (ii) as construções de tópico não são uma inovação dentro da comunidade, principalmente as de Tópico Pendente com Retomada, Tópico Cópia, Tópico Sujeito, Duplo Sujeito e Topicalização Selvagem, uma vez que o número de realizações

desses tópicos nos informantes da faixa 3 que permaneceram na comunidade aproxima-se do número de realizações dos da faixa 2 que saíram da comunidade (e, às vezes, o ultrapassa); (iii) o nível de escolaridade não interfere na realização das construções de tópico. Este último fato é confirmado ao se tomar como parâmetro as construções atestadas por Pontes (1986, 1987) do português brasileiro urbano na fala de pessoas escolarizadas, que também realizam os mesmos tipos de tópico encontrados nas comunidades em estudo; (iv) a faixa etária parece ter interferência nas realizações das construções de tópico.

Os dados encontrados em relação à faixa etária nessas comunidades trazem uma importante contribuição a outras pesquisas que analisam o português formal de peças de teatro ou de cartas pessoais, como as realizadas por Decat (1989) e por Araújo (2006), dentre outras. Essa contribuição se refere ao fato de terem sido encontradas construções de Tópico Sujeito, de Topicalização de Objeto Direto e de Topicalização Selvagem na fala de pessoas com 103 anos ou 107 anos de idade, como mostram os dados de Helvécia e de Cinzento, o que implica a realização dessas construções ter começado ao menos no final do século XIX ou início do século XX. Vejamos por quê. A coleta de dados nessas comunidades deve estar completando mais ou menos 20 anos, o que implica que essas pessoas de 103 e de 107 anos, no início do século XX, deveriam estar com a aquisição da linguagem consolidada. Essas construções de tópico deveriam estar sendo oralmente utilizadas e talvez tenham sido aí adquiridas por elas. Os exemplos abaixo dão uma indicação:

Informante com 103 anos:

(51) *Eu já morreu*[18] *quatro fi'.* (HV-19)

(52) *Tudo me sentiu má como esse fio, eu senti má!* (HV-19)

(53) *Meus neto d'eu conversa quando vem de longe!* (HV-19)

Informante com 107 anos:

(54) *Reza véa, pode num sabê*... (CZ-10)

Nessas realizações, é possível identificar construções com Tópico Sujeito (51-52), Topicalização Selvagem (53) e Topicalização de Objeto Direto (54).

Considerando esses dados hipoteticamente, temos evidências de que essas construções, atuais no português brasileiro, já estavam presentes na oralidade no final do século XIX ou no início do século XX sendo também possível que tenham surgido a partir da influência do contato do português com as línguas africanas. É preciso esclarecer, para os propósitos do trabalho, que, sendo as comunidades analisadas de afrodescendentes, principalmente a de Helvécia (a mais antiga na

[18] É preciso ressaltar que pessoas mais velhas nessa comunidade não realizam a concordância verbal na primeira pessoa do singular (cf. nota 7).

coleta de dados), defendemos a possibilidade de haver interferência do contato entre línguas na formação do português brasileiro. Concordamos com Lucchesi (1999a) quando trata do processo de transmissão irregular do português aos africanos que aqui chegaram sem o domínio dessa língua, o que pode ter desencadeado uma crioulização leve. Tal forma de crioulização, para o autor, não levaria a uma reestruturação gramatical completa da língua, mas poderia desencadear (no nosso entendimento) a reanálise em algumas áreas sintáticas, como a do sujeito, por exemplo. Desse modo, assumimos que, no processo de aprendizagem irregular da língua portuguesa, os africanos podem ter recorrido ao conhecimento de suas línguas de origem, o que pode ter implicado reanálise de algumas funções sintáticas em seu processo de desempenho linguístico da língua de superestrato. Temos evidência disso na fala das pessoas mais velhas da comunidade de Helvécia e Cinzento, como demonstrado acima, e que parece ter originado as atuais construções de tópico presentes no português brasileiro.

O cruzamento dos dados linguísticos com os dados sociais indica que as construções de tópico nessas comunidades estão ligadas à faixa etária mais velha. À primeira vista, esses dados poderiam fornecer evidências contrárias com relação às análises do português urbano moderno; entretanto, do que sabemos, até agora não se procedeu a um controle de cruzamentos de dados linguísticos e sociais no que se refere às construções de tópico no português urbano moderno, o que implica não haver informações suficientes que possam refutar ou endossar os resultados encontrados no português rural afro-brasileiro.

9.5 Considerações finais

A partir da análise realizada das construções de tópico no *corpus* do português rural afro-brasileiro, ressaltamos os seguintes pontos:

- as construções de tópico presentes nessas comunidades também estão presentes no português urbano (ou será o contrário?);
- parece haver uma tendência para maior número de realizações de tópico por pessoas mais velhas dentro das comunidades, o que reforça a hipótese da influência do contato entre línguas na formação do português brasileiro;
- os dados levam à evidência de que a realização oral de construções de Tópico Sujeito, Topicalização do Objeto Direto e Topicalização Selvagem pode ter começado já no início do século XX;
- apesar do levantamento numérico realizado, cruzando os tipos de tópico e os dados sociais, como permanência fora da comunidade, nível de escolarização e idade, reconhecemos que, quando se trata de construções de tópico,

a quantificação não fornece um quadro preciso, porque essas construções estão na interface do discurso com a sintaxe e, portanto, subordinadas ao contexto. Ou melhor, a análise das construções de tópico não se restringe apenas à quantidade de vezes que o falante recorre a elas, mas à sua existência, visto que a produção dessas construções revela os procedimentos sintáticos empreendidos pelos falantes em seu desempenho linguístico para atender aos requerimentos discursivos demandados pelo contexto.

Levando em consideração a semelhança entre as construções de tópico do português urbano com o português rural afro-brasileiro, surge um inquietante questionamento: afinal, o português brasileiro urbano influenciou ou sofreu influência do português rural afro-brasileiro? Em princípio, acreditamos que o segundo tenha influenciado o primeiro,[19] como defende Tarallo (1993b), para quem, o processo de migração do campo para a cidade favoreceu o intercâmbio linguístico e, consequentemente, a interferência dos falares rurais nos falares urbanos. Salientamos, entretanto, que pesquisas mais aprofundadas a esse respeito devem ser realizadas para responder a essa questão.

[19] Há toda uma discussão histórica a esse respeito, como o número de negros na população brasileira, a extensão de sua interferência linguística, os movimentos migratórios no Brasil, entre outros dados. Para mais informações a esse respeito, cf. Tarallo (1993b), Mattos e Silva (1998), Lucchesi (1999a), dentre outros.

10 A negação sentencial

*Rerisson Cavalcante**

Este capítulo trata da expressão da negação sentencial em uma variedade específica do português brasileiro (PB), o dialeto de comunidades rurais isoladas, constituídas majoritariamente por afrodescendentes.[1] No PB contemporâneo, coexistem três padrões de negação sentencial, de acordo com a posição em que a partícula negativa aparece em relação ao sintagma verbal (VP): um primeiro padrão, *pré-verbal* [Não VP], conforme (1); um segundo padrão, em que ocorrem dois marcadores, um em posição pré- e outro pós-verbal [Não VP não], conforme (2); e um terceiro padrão, *pós-verbal* [VP não], como em (3):

(1) **Não** sei o nome dela. [Não VP]

(2) **Não** sei o nome dela **não**. [Não VP não]

(3) Sei o nome dela **não**. [VP não]

O objetivo principal deste capítulo é a identificação de aspectos linguísticos e discursivos, além de sociais, que condicionam o uso da negação no dialeto examinado, bem como de restrições sintáticas à ocorrência das variantes. O texto está organizado da seguinte forma: na seção 10.1, apresento os aspectos metodológicos da pesquisa; na seção 10.2, a descrição e análise quantitativa dos dados em função dos aspectos linguísticos examinados; e na seção 10.3, do ponto de vista dos aspectos sociais. Por fim, na seção 10.4, apresento as conclusões.

10.1 Metodologia

Esta pesquisa foi realizada de acordo com a metodologia de base estatística desenvolvida pela Sociolinguística Variacionista (LABOV, 1983 [1972a]). O *corpus* utilizado é constituído por entrevistas informais com membros de três comunidades

* Meus agradecimentos a Ilza Ribeiro e Jacyra Mota, respectivamente orientadora e co-orientadora da dissertação de Mestrado de que esse trabalho se origina.

[1] Uma análise mais ampla da negação sentencial no português brasileiro se encontra em Cavalcante (2007).

afro-brasileiras do interior do Estado da Bahia: Cinzento (no Município de Planalto), Sapé (no Município de Valença) e Barra e Bananal (no Município de Rio de Contas).[2] O recorte do *corpus* é constituído por dezoito entrevistas, seis de cada comunidade, realizadas com informantes naturais da localidade. Os informantes são distribuídos por sexo e por três faixas etárias, a saber: a faixa 1, com informantes entre 20 e 40 anos; a faixa 2, entre 41 e 60 anos; e a faixa 3, com mais de 60 anos.[3]

10.2 Descrição dos dados: aspectos linguísticos

Em um total de 2.026 sentenças levantadas do *corpus*, a negativa pré-verbal [Não VP] ocorre com uma frequência de 66%. Já a estrutura com dois marcadores negativos, [Não VP não], corresponde a 28% dos dados, e a negativa pós-verbal [VP não] a apenas 6%, conforme números da Tabela 1.[4]

Tabela 1 - Distribuição das construções negativas no *corpus*

	[NÃO VP]	[NÃO VP NÃO]	[VP NÃO]	TOTAL
Ocorrências	1343	568	115	2026
Frequência	66%	28%	6%	100%

Resultados de trabalhos sobre outras variedades do PB (RONCARATI, 1996; FURTADO DA CUNHA, 1996; CAMARGOS, 2000; ALKMIM, 2002; SOUZA, 2004 confirmam que a negativa pré-verbal é a estratégia que ocorre majoritariamente. Entretanto, a comparação dos números de ocorrências e porcentagens mostra que o perfil da formação demográfica e histórica das comunidades se reflete na difusão das variantes.

---

2 Para uma maior informação sobre as comunidades analisadas, ver o capítulo 2 deste livro.

3 A metodologia de constituição dos *corpora* das localidades estudadas é apresentada no capítulo 5.

4 Os advérbios e quantificadores negativos do tipo ***nada***, ***ninguém***, ***nenhum(a)*** e ***nunca*** não foram tratados como marcadores de negação, uma vez que possuem o traço negativo como secundário em relação a outras funções que desempenham, além de manifestarem outras propriedades sintáticas distintas do NÃO. Dessa forma, sentenças como (i) e (ii), abaixo, foram consideradas como casos de [Não VP] e [Não VP não], respectivamente.

(i) [Não VP]
 (a) (...) sei lá, **num** tava sentino quase **nada** também. (RC-04)
 (b) poque **num** conhecia **ninguém**. (RC-24)
 (c) **num** teve confusão **nenhuma**. (RC-05)

(ii) [Não VP não]
 (a) Ela levô um dia lá no hospital, mas **num** teve **nada não**. (RC-13)
 (b) **num** vi **nada não**. (RC-24)

A negação pré-verbal [Não VP] apresenta frequências maiores em dialetos de comunidades com perfil demográfico mais complexo e sem um histórico de forte (e prolongado) contato linguístico e sem isolamento, como as capitais Fortaleza, Natal, Belo Horizonte e a cidade mineira de Mariana.[5] A frequência, por outro lado, cai nas localidades isoladas e em que o contato linguístico foi predominante e prolongado, como Pombal, Helvécia[6] e as três comunidades investigadas neste trabalho.

No primeiro caso, os percentuais da negativa pré-verbal oscilam entre 77% e 88%, enquanto nas outras comunidades esses números ficam entre 64% e 66%, como pode ser visto na Tabela 2.

Tabela 2 - Frequência de [Não VP] *versus* perfil das localidades

PERFIL DAS COMUNIDADES	COMUNIDADES	[NÃO VP]
Comunidadesnão marcadas etnicamente	Natal	88,6%
	Fortaleza	77%
	Belo Horizonte	70%
	Mariana	77,1%
Comunidades de origem afro estudadas na pesquisa	Cinzento, Sapé, Rio de Contas	66%
Outras comunidades de origem afro	Helvécia	66,9%
	Pombal	64,2%

A seguir, apresento os resultados quantitativos de cada um dos fatores investigados nos dados das comunidades aqui analisadas.

10.2.1 Negação sentencial e tipo de frase

Nesta primeira variável explanatória, examino a distribuição das sentenças negativas pelo tipo de frase em que ocorrem, considerando três possibilidades: se a oração negativa ocorre em um contexto de *resposta*; em um contexto de *não resposta* ou como *pergunta*, conforme descrito a seguir:

(i) frase dada como **resposta a uma pergunta direta**, entendida como resposta a uma pergunta polar (*yes/no-question*)

(4) *Respostas a pergunta polar*

(a) E ela costuma ir?

Num *gosta muito de ir im festa* ***não***...(SP-04)

---

[5] Os dados de Fortaleza são de Roncarati (1996); os de Natal são de Furtado da Cunha (1996); os de Belo Horizonte, de Camargos (2000), e os de Mariana, de Alkmim (2002).

[6] Os dados de Pombal são de Alkmim (2002), e os de Helvécia, de Souza (2004).

(b) E demora muito pa descê, não? Ô vai esperano secá, não?

Não. *Demora **não**...* (SP-04)

(ii) frase usada como **pergunta**, aí enquadrados casos de perguntas polares e perguntas-QU[7]

(5) *Perguntas*

(a) Eu vi, vai, vai, vai... E o rapaz tamém, *é noivo **não**?* (SP-09)

(b) Tem... ***Num** viu onte ali na venda o tanto de mulhé que tinha, **não**?* (SP-01)

(iii) e frases que não funcionam como resposta direta nem como pergunta, denominadas genericamente de **não resposta**.

(6) *Não resposta*

(a) Aí ela pegô esse carajé, por causa de me dá todim p'eu comê, pegô e ***num** dividiu*. (SP-05)

(b) Aí nós... aí agora nós ficô na praça, assim camiano pra lá e prá cá, *nós **num** parava **não**...* (CZ-01)

Na delimitação dessa variável, foi importante a hipótese de que o contexto de respostas curtas e/ou de respostas diretas seria um favorecedor do uso da negativa pós-verbal [VP não]. Tal ideia é apontada, mesmo impressionisticamente, por muitos pesquisadores (RONCARATI, 1996; FURTADO DA CUNHA, 1996, 2001). Furtado da Cunha (1996) defende que a "negação dupla" é usada para recusar oferta ou sugestão, ou para rejeitar uma asserção previamente mencionada ou pressuposta pelos interlocutores. Desse modo, a partícula de negação pós-verbal introduziria informação preposicional nova, "mas para negar informação já presente no discurso ou ao menos compartilhada por falante e ouvinte pelo conhecimento de mundo" (p.173).

A variável tipo de frase foi selecionada pelo VARBRUL como a mais importante no condicionamento da realização da negação sentencial. Os resultados percentuais e os pesos relativos, apresentados na Tabela 3, mostram que as negativas em que aparece uma partícula de negação em posição pós-verbal, ou seja, [Não VP não] e [VP não], se comportam de maneira oposta a [Não VP] com relação a essa variável.

---

[7] Perguntas-QU são aquelas em que aparece um pronome interrogativo, que funciona como uma variável a que a resposta deve fornecer uma especificação. Uma *pergunta polar* não apresenta pronome interrogativo e exige a confirmação ou refutação do conteúdo proposicional da sentença, em geral, por partículas assertivas com *sim* e *não*:

(i) — Quem você viu com Maria na festa?
— (Eu vi) João.
[Quem = João]

(ii) — Você viu João na festa com Maria?
— Sim / Não / Vi.

Tabela 3 - Tipo de frase e tipo de negativa sentencial

TIPO DE FRASE	[NÃO VP]		[NÃO VP NÃO]		[VP NÃO]		TOTAL
	%	P.r.	%	P.r.	%	P.r.	
Não resposta	73	**.55**	24	.46	2	.43	1687
Pergunta	50	.46	40	**.59**	10	**.62**	52
Resposta a *yes/no question*	28	.24	47	**.68**	25	**.83**	287

Nível de significância: .000 para [Não VP], .008 para [Não VP não], .005 para [VP não].

Enquanto a variante pré-verbal [Não VP] é favorecida pelo contexto de não resposta e desfavorecida por respostas diretas e perguntas, as duas outras variantes [Não VP não] e [VP não] são favorecidas pelo contexto de respostas, principalmente, e pelo de perguntas, sendo desfavorecidas pelo de não respostas. Esses resultados são confirmados pelo trabalho de Souza (2004), com dados de Helvécia, outra comunidade isolada de origem africana do interior da Bahia (Município de Nova Viçosa).

Apesar de o contexto de perguntas aparecer como favorecedor das negativas com ***não*** final, estas mesmas negativas não ocorrem no *corpus* em perguntas-QU, mas somente em perguntas polares, conforme exemplos em (7).

(7) (a) Tu já foi **não**? (CZ-09)

(b) Ah, se ocê... Tu faz jarro dento de casa **não**? (CZ-09)

(c) Ah! O povo de Lula que se exibe?

É. Os eleitô dele que se exibe. Parece que vai até dá um bom... presidente, né? Tá dano aí comida aí aos... *Tu viu na televisão **não***? (SP-01)

A ausência das negativas [Não VP não] e [VP não] em perguntas-QU no *corpus* pode indicar que essas construções são agramaticais nesse contexto. A comparação dos exemplos em (9-10) com (8) mostra que elas têm, no mínimo, uma gramaticalidade duvidosa.

(8) (a) O que foi que ele **não** fez?

(b) Por que ele **não** saiu de casa?

(c) Que horas você **não** está ocupado?

(9) (a) ?*O que foi que ele **não** fez **não**?

(b) ?*Por que ele **não** saiu de casa **não**?

(c) ?*Que horas você **não** está ocupado **não**?

(10) (a) ?*O que (foi que) ele fez **não**?

(b) ?*Por que ele saiu de casa **não**?

(c) ?*Que horas você está ocupado **não**?

O Quadro 1 resume a ação da variável tipo de frase.

TIPO DE NEGATIVA	FAVORECIDA POR	DESFAVORECIDA POR	AGRAMATICAL EM
[Não VP]	não resposta	perguntas e respostas diretas	——
[Não VP não] [VP não]	perguntas e respostas diretas	não resposta	perguntas-QU

Quadro 1 - Resumo da atuação da variável *tipo de frase* sobre as negativas

De um modo geral, os resultados confirmam que [Não VP] é uma estrutura não marcada, usada em contextos de negação de uma declaração; já [Não VP não] e, principalmente, [VP não] estão ligadas a contextos de respostas diretas, em que se expressa não apenas a negação da própria sentença, mas de pressupostos ativados, por um interlocutor, numa situação de diálogo (GIVÓN, 1984; RONCARATI, 1996).

10.2.2 Negação sentencial e tipo de oração

A variável tipo de oração leva em conta o tipo ou a função da oração em que a negação aparece. Os resultados do VARBRUL apontam esse grupo de fatores como o segundo mais importante no condicionamento da negação, pois foi o segundo fator selecionado para [Não VP não] e [VP não] e o terceiro para [Não VP]. As orações foram classificadas em oito tipos, descritos abaixo:

(11) *Absoluta*

(a) **Num** sei **não**. (SP-01)

(b) **Num** saiu nenhum ainda **não**. (SP-01)

(c) Hum... Era muntcho boa **não**. (SP-01)

Principal

(d) Aí ficô lá, ***num*** *sei* se interno... (RC-04)

(e) Ieu num sei... do... *ieu* ***num*** *sei* do quê qu'ele tá trabaiano lá ***não***... (RC-04)

(f) que é umas rosas feita assim... num sei se é de papel, *sei* que que é ***não***...

Coordenada "livre"[8]

(g) ... *a gente* ***num*** *vamo de carro*, a gente vamo... vamo montado... (RC-05)

(h) ... e dexo aí mesmo... ***num*** *levo pra lá* ***não***. (CZ-05)

(i) *Guenta* ***não***! A terra aqui é muito ININT (RC-08)

Coordenada introduzida por conjunção

(j) É, tem gente que continua, moço! *Mas aí é* ***num*** *tá mais estimano a vida*, né? (RC-08)

[8] Ao longo do texto, por questões de comodidade e de facilidade mnemônica, a *coordenada não introduzida por conjunção* será chamada de *coordenada livre*.

(k) ... num quero ir. *Mas eu **num** qué, **não***. (RC-05)

(l) Ele nunca que chega, assim, de uma vez, *e estraga **não*** (CZ-04)

Substantiva

(m) A gente ficô assim... pensano *que **num** ia voltá estudá*, né? (CZ-01)

(n) É, eu acredito *que a metade **num** volta **não***... (CZ-01)

Adverbial

(o) Quando eu **num** tava aqui, minha irmã cozinhava... (CZ-08)

(p) Ensina. Se a pessoa **num** entendê uma coisa... (SP-01)

Relativa

(q) ... tem um período de tempo *qu'eu **num** viajei **não***. (CZ-03)

(r) Chega lá compra o ingresso *que **num** tem* e só que já era mais, né? (CZ-08)

Causal / explicativa[9]

(s) ... *poque tamém o terreno **num** dá* pra fazê direto, né? (RC-26)

(t) Vai vê que já [criaro] estrada pra lá, *poque **num** tinha estrada de carro po Paramirim não*. (RC-26)

(u) Aí ela chegava na sala com a cara feia, *que **num** gostava*. (SP-01)

As porcentagens e os pesos relativos apresentados na Tabela 4 mostram que a estrutura [Não VP] é favorecida por todos os tipos de sentenças encaixadas, e [Não VP não] por todas as matrizes. Já [VP não] é favorecida apenas por dois tipos de matrizes.

Tabela 4 - Tipo de oração e tipo de negativa sentencial

TIPO DE ORAÇÃO	[NÃO VP]		[NÃO VP NÃO]		[VP NÃO]		TOTAL
	%	P.r.	%	P.r.	%	P.r.	
Absoluta	26	.27	37	**.52**	36	**.89**	118
Principal	69	.48	30	**.53**	1	.22	319
Coordenada assindética	60	.48	32	**.53**	7	**.64**	913
Coordenada sindética	66	.45	33	**.58**	1	.20	249
Substantiva	77	**.60**	23	.45	——	——	86
Adverbial	100	——	——	——	——	——	112
Causal / explicativa	85	**.69**	15	.33	——	——	154
Relativa	90	**.81**	10	.21	——	——	73

Nível de significância: .000 para [Não VP], .008 para [Não VP não], .005 para [VP não].

[9] Ao invés de serem classificadas, respectivamente, como *coordenadas introduzidas por conjunções* e como *encaixadas adverbiais*, as orações explicativas e causais foram agrupadas em uma categoria à parte. Essa decisão teve o objetivo de prevenir o enviesamento dos dados que poderia ocorrer, uma vez que os critérios para distinguir os dois tipos de oração não são nítidos em português, principalmente com dados de fala.

Negativa pré-verbal [Não VP]

A negativa pré-verbal ocorre majoritariamente nas orações encaixadas: é categórica nas sentenças *adverbiais* (100% dos dados) e tem frequências de 90%, 85% e 77% nas *relativas*, *causais* e *substantivas*, respectivamente. Os valores dos pesos relativos são: de .81 em relativas, de .69 em causais e de .60 em substantivas.

Entre as sentenças matrizes, por outro lado, [Não VP] sofre um leve desfavorecimento por parte das *coordenadas* e das *principais*, mas parece ser fortemente desfavorecida pelas orações *absolutas*. Os números dos pesos relativos são de .48 em *principais* e *coordenadas livres* e de .45 nas *coordenadas introduzidas por conjunções*. Já nas *absolutas*, o peso relativo de [Não VP] é de apenas .27.

Negativa pré- e pós-verbal [Não VP não]

A negativa [Não VP não] também se comporta de modo oposto à pré-verbal. É desfavorecida por sentenças encaixadas, como mostram os pesos relativos de .21 para as *relativas*, de .33 para as *causais* e de .45 para as *substantivas*. Além disso, a construção não foi documentada no *corpus* para as orações *adverbiais*.

As sentenças matrizes, por outro lado, aparecem como um contexto de favorecimento da negativa [Não VP não], que obteve peso relativo de .58 nas *coordenadas introduzidas por conjunção*, de .53 nas *principais* e nas *coordenadas livres* e de .52 nas *absolutas*.

Em resumo, sentenças encaixadas favorecem [Não VP] e desfavorecem [Não VP não], enquanto matrizes desfavorecem [Não VP] e favorecem [Não VP não].

Negativa pós-verbal [VP não]

A negativa pós-verbal [VP não] se comporta de maneira apenas parcialmente semelhante a [Não VP não]. Apesar de também ser desfavorecida por sentenças encaixadas, [VP não] não é favorecida por todos os tipos de matrizes, como [Não VP não] o é. A negativa final recebe um incremento por parte das *coordenadas livres* e, principalmente, pelas *absolutas*, como mostram os pesos relativos de .64 e .89, respectivamente. Já as *principais* e as *coordenadas iniciadas por conjunção* aparecem desfavorecendo fortemente a variante, conforme os pesos de .22 e .20.

O tipo de oração indica ainda a existência de contextos sintáticos em que [VP não] não ocorre no *corpus* examinado. A construção não é encontrada em frases encaixadas de nenhum tipo, sejam *subjetivas*, *adverbiais*, *causais* ou *relativas*. Esse é mais um aspecto que distingue [VP não] de [Não VP não], pois, com exceção das *adverbiais*, a última é somente desfavorecida nesses contextos, mas ainda assim ocorre com frequências que variam entre 10% e 23%. O Quadro 2 resume a atuação dessa variável sobre a negação.

TIPO DE NEGATIVA	FAVORECIDA POR	DESFAVORECIDA POR	AGRAMATICAL EM	CATEGÓRICA EM
[Não VP]	encaixadas	matrizes	——	encaixadas adverbiais
[Não VP não]	todas as matrizes	encaixadas	encaixadas adverbiais	——
[VP não]	matrizes absolutas e coordenadas livres	matrizes principais e iniciadas por conjunção	todas as encaixadas	——

Quadro 2 - Resumo da atuação da variável *tipo de oração*

10.2.3 Negação sentencial e realização dos argumentos verbais: complementos verbais

Nesta e na próxima subseção, apresento os resultados das variáveis ligadas ao tipo de realização dos argumentos verbais. As variáveis *tipo de sujeito* e *tipo de complemento verbal* buscam contrapor a realização *versus* a não realização dos argumentos verbais, entendendo que a omissão desses está ligada à possibilidade de já terem sido mencionados anteriormente. Além disso, essas variáveis levam em consideração também a posição em que o argumento é realizado e a natureza da não realização. Dessa forma, com relação ao complemento verbal, os dados foram assim classificados:

(i) ***complemento realizado in situ***, quando o constituinte em função de objeto do verbo está em posição pós-verbal.

(12) (a) Eu **num** tenho *dinhêro*, (SP-12)

(b) Eu nem sei... **num** sei a idade dele **não**! (RC-13)

(c) Não. Dá muito trabalho **não**! (SP-06)

(ii) ***complemento topicalizado ou anteposto***, quando o objeto do verbo aparece movido para a posição inicial da sentença ou simplesmente para uma posição à direita de sujeito e verbo.[10]

(13) (a) *Na de Bom Jesus mesmo*, nós **num** foi. (CZ-08)

(b) ... mas *com o arame* eu **num** preocupo **não**. (CZ-06)

(c) ... *dinhêro* eu **num** tinha. (CZ-11)

(iii) ***complemento não realizado***, que se trata do objeto que ocorre nulo ou por já ter sido mencionado no discurso ou por ser recuperável pelo contexto, representado nos exemplos por Ø.

[10] A segunda possibilidade, contudo, é citada apenas como uma possibilidade lógica, uma vez que não foram registrados casos de ordem SOV no *corpus*.

(14) *Complemento nulo*

(a) Ah! Rita foi lá em cima levá um pessoar, aí... tinha uns rapaz lá, *mas eu **num** conheci Ø*. (RC-24)

(b) E porque que lá se chama Rapa Tição, cê sabe?

É... *Eu **num** sei Ø **não***... (SP-12)

(c) *Vai Ø **não**.... Vai Ø **não***. (SP-12)

(iv) ***complemento não projetado***, que corresponde aos casos de verbos intransitivos, que não possuem complemento a ser projetado, ou de transitivos usados intransitivamente.[11]

(15) *Complemento não projetado*

(a) Trabaio. Agora que eu num tava... *qui eu **num** trabaiei* Δ poque eu tava doenti... (RC-04)

(b) ... **num** existia isso aí tamém **não**. (RC-05)

(c) Aí ININT o osso de cá é mais duro... Eu corro Δ **não**. (SP-01)

A variável *tipo de complemento* foi selecionada pelo VARBRUL como o quinto fator mais importante para a realização de [Não VP] e como o sexto para a [Não VP não], mas não foi selecionada para [VP não]. Os valores da Tabela 5 mostram que as negativas [Não VP] e [Não VP não] apresentam igualmente uma distribuição oposta.

Tabela 5 - Realização do complemento verbal e tipo de negativa sentencial

TIPO DE COMPLEMENTO VERBAL	[NÃO VP]		[NÃO VP NÃO]		[VP NÃO]		TOTAL
	%	P.r.	%	P.r.	%	P.r.	
Realizado *in situ*	72	**.54**	25	.47	3	—	1093
Topicalizado	80	**.70**	20	.36	—	—	41
Não-realizado / nulo (Ø)	57	.44	32	**.54**	11	—	705
Não-projetado (verbo intransitivo)	62	.42	33	**.56**	5	—	168

Nível de significância: .000 para [Não VP], .008 para [Não VP não].

A negativa [Não VP] é favorecida pela realização lexical do complemento, seja *in situ* ou *topicalizado*, como mostram os pesos relativos de .54 e .70, respectivamente, e é desfavorecida pelo complemento *nulo* ou complemento *não projetado*, com pesos de .44 e .42.

[11] O símbolo Δ representa a inexistência do complemento, em oposição a Ø, que representa a sua omissão.

Por outro lado, complementos *nulos* e *não projetados* favorecem levemente [Não VP não], com peso de .54 e .56, respectivamente. Já os complementos realizados desfavorecem essa variante, principalmente os *topicalizados*, que têm peso de .36 enquanto os *in situ* têm .47.

Embora tenha sido descartado para [VP não], esse fator traz uma importante informação sobre essa construção, pois aponta para um possível contexto de restrição à sua ocorrência. Nos dados das três comunidades, não há casos de [VP não] com o complemento *topicalizado*, o que pode indicar mais um contexto de agramaticalidade desse tipo de negativa, como mostram os exemplos em (16), criados como teste de gramaticalidade:[12]

(16) (a) ??*Esse livro* eu comprei **não**.

(b) ??*Com o arame* eu preocupo **não**.

(c) ??*Feijão* eu gosto **não**.

(d) ??*Show de Ivete* eu vi **não**.

(d) ??*Esse ladrão* o policial prendeu **não**.

10.2.4 Negação sentencial e realização dos argumentos verbais: sujeito

Com relação à variável tipo de sujeito, foram consideradas quatro possibilidades:

(i) ***sujeito realizado anteposto***, que corresponde aos dados em que o sintagma que exerce a função de sujeito está realizado foneticamente em posição pré-verbal.

(17) Sujeito realizado anteposto ao verbo

(a) ... pode dizê que *eu* **num** conheço ga... (CZ-26)

(b) *Eu* **num** fiquei muito tempo **não**. (CZ-26)

(c) ... *você* vai com a menina hoje **não**. (SP-05)

(ii) ***sujeito realizado posposto***, ou seja, sujeito realizado foneticamente em posição à direita do verbo.

(18) Sujeito realizado posposto ao verbo

(a) ... **num** vai *duas pessoa*... (SP-01)

(b) **Num** saiu *nenhum* ainda **não**. (SP-01)

(iii) ***sujeito nulo referencial***, que é o caso em que o sujeito temático é omitido em razão da sua recuperabilidade pela morfologia e pelo contexto linguístico.

[12] Versões dessas mesmas frases, com a presença do marcador negativo pré-verbal, estão presentes no *corpus*.

(19) Sujeito nulo referencial (Ø)

(a) ... *quano Ø* ***num*** *queria estudá*, eu saía. (SP-01)

(b) Ø **Num** sei **não**. (SP-01)

(c) Ah, num fala não, que... *Ø Fala* ***não***. (SP-05)

(iv) e ***sujeito nulo expletivo***, que corresponde aos casos de construções existenciais ou com verbos que indicam fenômenos atmosféricos, que não possuem sujeito temático.

(20) Sujeito nulo expletivo (Δ)

(a) *Se Δ* ***num*** *chovê*, num nasce. (SP-09)

(b) Mas Δ **num** tinha **não**. (CZ-26)

(c) Δ Tem **não**. (CZ-11)

Esse fator foi selecionado como o terceiro mais relevante para [VP não], sendo descartado pelo VARBRUL para os outros padrões de negação.

De um modo geral, com relação ao tipo de realização do sujeito, as negativas se comportam de modo paralelo[13] à realização do complemento verbal, com [VP não] sendo favorecida por sujeitos nulos tanto *referenciais* quanto *expletivos*, que apresentam, ambos, o peso relativo de .66, e é desfavorecida pela sua realização lexical em posição pré-verbal, que tem peso de .31, como se pode verificar na Tabela 6.

Tabela 6 - Realização do sujeito e tipo de negativa sentencial

TIPO DE REALIZAÇÃO DO SUJEITO	[NÃO VP]		[NÃO VP NÃO]		[VP NÃO]		TOTAL
	%	P.r.	%	P.r.	%	P.r.	
Realizado	70	—-	28	—	2	.31	879
Realizado posposto	69	—-	31	—	—	—	29
Nulo referencial (Ø)	61	—	30	—	9	**.66**	875
Nulo expletivo (Δ)	68	—	22	—	10	**.66**	213

Nível de significância: .005 para [VP não].

A realização do sujeito fora de sua posição "típica", pré-verbal, parece, a princípio, ser um contexto de restrição à ocorrência de [VP não]: assim como não há exemplos no *corpus* de [VP não] com complemento topicalizado, não há dados dessa construção com sujeito posposto.

A posposição parece ser possível pelo menos com verbos inacusativos (e parece até mesmo bastante natural), principalmente quando o sujeito é um Item de Polaridade Negativa (*nada, ninguém, nenhum*), conforme exemplos em (21).

---

[13] A comparação não se dá diretamente com relação a [VP não], uma vez que o tipo de complemento não foi selecionado para essa negativa, mas entre [VP não] e [Não VP não], estruturas em que há um marcador negativo em posição final, *versus* [Não VP].

(21) (a) Veio *ninguém* **não**.

(b) Aconteceu *nada* **não**

(c) ? Chegou *o menino* **não**

Dois fatores podem ter contribuído para a ausência de dados como em (21) no *corpus*: (i) a própria baixa ocorrência de sujeitos pospostos de qualquer tipo, que se restringiu a 19 casos; (ii) e a não ocorrência de Itens de Polaridade Negativa como sujeitos pospostos mesmo com os outros tipos de negativas, [Neg V] e [Neg V neg].

O Quadro 3 resume a atuação das variáveis tipo de sujeito e tipo de complemento.

TIPO DE NEGATIVA	FAVORECIDA POR	DESFAVORECIDA POR	AGRAMATICAL EM
[Não VP]	complemento realizado *in situ* ou topicalizado	complemento nulo e não projetado	——
[Não VP não]	complemento nulo e não projetado	complemento realizado *in situ* ou topicalizado	——
[VP não]	sujeito nulo referencial e expletivo	sujeito realizado	complemento topicalizado

Quadro 3 - Resumo da atuação das variáveis *tipo de sujeito* e *tipo de complemento*

10.2.5 Negação sentencial e realização fonética da partícula pré-verbal

Outra variável linguística examinada foi a realização fonética da partícula pré-verbal de negação. Nessa posição, o marcador negativo pode ser pronunciado como ***não*** e como ***num***.[14] Os números da Tabela 7 mostram que a forma ***num*** é quase categórica nessa posição na variedade estudada, ocorrendo com uma média de 96% dos dados.

Tabela 7 - Realização fonética da partícula negativa pré-verbal nas três comunidades

REALIZAÇÃO FONÉTICA	LOCALIDADES CINZENTO		SAPÉ		RIO DE CONTAS		TOTAL	
num	754	94,36%	620	97,79%	428	99,07%	1802	96,62%
não	45	5,63%	14	2,2%	4	0,92%	63	3,37%

[14] O ***não*** pode ocorrer, também, como uma simples consoante nasal [n] diante das formas verbais "*é*" e "*era*", como em (i) e (ii). Esses casos foram excluídos da análise, uma vez que atendem a um condicionamento mais lexical do que sintático ou fonético-fonológico: aparecem apenas com essas duas formas do verbo *ser*. Ocorrências como (iii) são agramaticais.
(i) **Né** não!
(ii) **N'** era nada não.
(ii) * N'empata não.

Os resultados de outros trabalhos mostram que a forma ***num*** é, provavelmente, a forma vernácula do PB. É a mais recorrente em vários dialetos documentados, e sua frequência ainda aumenta à medida que a variedade demonstra menos influência da norma escrita e da escolarização: tem frequência de 59% em Belo Horizonte (RAMOS, 2002), 86% em João Pessoa (MATA, 1999) e 97% nos dados das comunidades rurais da Bahia aqui analisados.

Essa variável foi selecionada como a segunda mais relevante para [Não VP] e como a quarta para [Não VP não].[15] As negativas [Não VP] e [Não VP não] voltam a apresentar comportamentos opostos. Os números da Tabela 8 mostram que a forma ***não*** favorece fortemente a negação pré-verbal [Não VP] e desfavorece [Não VP não], com os pesos relativos bastante expressivos de .72 e . 29, respectivamente.

Tabela 8 - Realização fonética do NÃO pré-verbal e tipo de negativa sentencial

FORMA FONÉTICA	[NÃO VP]		[NÃO VP NÃO]		TOTAL
	%	P.r.	%	P.r.	
não	87	**.72**	13	.29	63
num	71	.49	29	**.51**	1802

Nível de significância: .000 para [Não VP], 008 para [Não VP não].

As duas construções, aparentemente, não se mostram tão sensíveis à forma ***num***, como se pode perceber pela proximidade dos pesos relativos de .50. A análise de cada construção separadamente, porém, indica um desfavorecimento de [Não VP] com a forma ***num***, já que a frequência da construção cai de 87% com ***não*** para 71%, o mesmo ocorrendo com o peso relativo, que vai de .72 para .49. Por outro lado, há um favorecimento de [Não VP não] pela forma ***num***, pois há um aumento da frequência de 13% com ***não*** para 29% com ***num*** e do peso relativo de .29 para .51.

10.3 Atuação dos aspectos sociais

Quatro variáveis sociais foram investigadas nesta pesquisa: (i) faixa etária dos informantes; (ii) sexo; (iii) nível de escolaridade e (iv) estada fora da comunidade por um período mínimo de seis meses. A única variável selecionada pelo VARBRUL como relevante no condicionamento da negação foi a estada fora da comunidade, que foi considerado o quarto aspecto mais importante para a realização de [Não VP] e o segundo para [Não VP não], mas foi descartado para [VP não].

[15] Note-se que esse grupo de fatores não se aplica a [VP não], uma vez que, nessa construção, inexiste um marcador pré-verbal.

Esta variável está relacionada com a importância que o contato com outros dialetos pode ter para a implementação de uma dada variante linguística. Avalia, assim, se há alterações no uso das negativas em função de o falante ter permanecido por algum tempo em contato direto com outra(s) variedade(s) linguística(s), principalmente urbana(s).

Os números percentuais e os pesos relativos indicam que a realização da variante pré-verbal [Não VP] aumenta entre aqueles que já passaram seis meses ou mais fora da comunidade e diminui entre os que nunca saíram, como mostra a Tabela 9. O oposto ocorre com [Não VP não]. Em outras palavras, a ocorrência da variante [Não VP] é favorecida pelo deslocamento e desfavorecida pela permanência na localidade; e [Não VP não] é desfavorecida pelo deslocamento da localidade e favorecida pela permanência.

Tabela 9 - Deslocamento da comunidade e tipo de negativa sentencial

ESTADA FORA DA COMUNIDADE	[NÃO VP]		[NÃO VP NÃO]		[VP NÃO]		TOTAL
	%	P.r.	%	P.r.	%	P.r.	
Sim	73	**.57**	22	.43	5	—	947
Não	60	.44	34	**.56**	6	—	1079

Nível de significância: .000 para [Não VP] e .008 para [Não VP não].

As variáveis *sexo* e *nível de escolaridade*, embora não tenham sido selecionadas pelo VARBRUL como relevantes para a realização da negação, apresentam, com relação aos números percentuais, um perfil semelhante ao da atuação do deslocamento. Ou seja: a variante [Não VP], considerada padrão, tem um pequeno aumento da frequência com os fatores que apontam para um maior contato com outras variedades linguísticas: o sexo masculino[16] (5% de aumento) e o contato com a escolarização (aumento de 11%).[17] Por outro lado, a variante [Não VP não] é favorecida pelos fatores que indicam um perfil conservador: o sexo feminino e a não escolarização. Para [VP não], sexo e escolaridade não apresentam variação percentual significativa. Quanto à variável faixa etária, os números não apontam para a existência de um processo de mudança em progresso, mas de variação estável, como pode ser visto na Figura 1.

[16] Os homens, por saírem com mais frequência dessas comunidades, por questões de trabalho e terem, assim, maior contato com outras variedades linguísticas, costumam estar mais próximos do padrão, como já mostraram outros estudos realizados no âmbito do Projeto *Vertentes* (LUCCHESI, 2002a).

[17] Os informantes estão distribuídos, quanto à escolaridade, em dois grupos: (i) os que tiveram algum contato com a escolarização, considerados *semialfabetizados* e (ii) os que não tiveram contato algum, sendo *analfabetos*. Esse perfil se deve à dificuldade de acesso à educação nessas comunidades, o que minimiza, em verdade, a força desse fator sobre a comunidade e pode ser o motivo para ele não ter sido selecionado pelo VARBRUL.

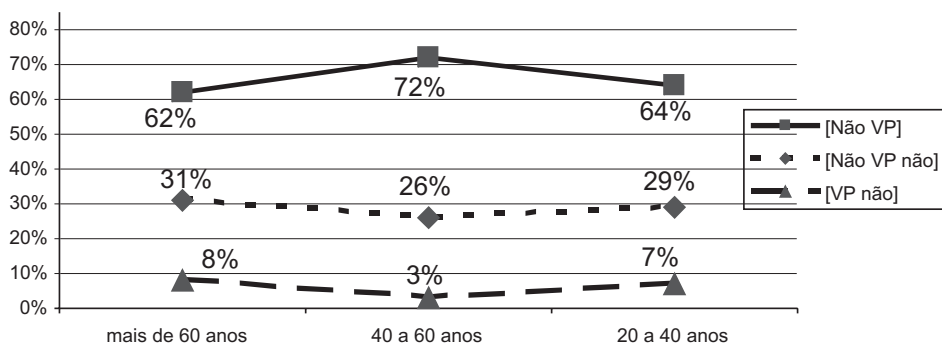



Figura 1 – Faixa etária e tipo de negativa sentencional

O Quadro 4 resume os resultados da atuação das variáveis sociais sobre a negação.

TIPO DE NEGATIVA	DESLOCAMENTO	SEXO	ESCOLARIDADE	FAIXA ETÁRIA
[Não VP]	Deslocamento	Masculino	Semianalfabetos	Variação estável
[Não VP não]	Permanência	Feminino	Analfabetos	
[VP não]	——	——	——	

Quadro 4 - Resumo da atuação das variáveis sociais sobre as negativas

10.4 Conclusões

Os resultados apresentados neste capítulo mostram que o uso das negativas sentenciais é fortemente condicionado por fatores linguísticos e discursivos. De um modo geral, foi possível confirmar a intuição de que o uso de [Não VP não] e [VP não] atende, prioritariamente, a uma função de negação de pressuposições. Isso se evidencia no favorecimento dessas variantes tanto pelos contextos de *perguntas* e *respostas diretas*, quanto pelo *apagamento dos argumentos verbais* e pelas *orações matrizes*, ao passo que a negativa [Não VP], considerada discursivamente neutra, é favorecida pelos contextos de *não resposta*, por *orações encaixadas* e pela *realização lexical dos argumentos verbais*.

Outro resultado importante alcançado pela pesquisa diz respeito à identificação de contextos sintáticos em que algumas variantes são bloqueadas, ou seja, situações de agramaticalidade. Como apontado, a negação [VP não] se mostra agramatical em três contextos: (i) em orações encaixadas de qualquer tipo; (ii) em orações com complementos topicalizados e (iii) em interrogativas-QU (a negativa [Não VP não] também é agramatical nesse último contexto).

Testes de gramaticalidade mostram que essas restrições não são válidas apenas para o dialeto examinado, mas provavelmente para todo o PB. No entanto, a

identificação das razões para a impossibilidade da ocorrência de [VP não] nesses contextos requer uma análise formal da gramática do PB que foge aos limites deste texto.[18]

Com relação aos aspectos sociais, as variáveis que indicam um maior contato com outras variedades linguísticas apontam um leve favorecimento da variante considerada padrão [Não VP] e um desfavorecimento da variante inovadora [Não VP não]. Por outro lado, [VP não] não é afetada por fatores sociais de nenhum tipo. Seu condicionamento é exclusivamente linguístico. Com relação à projeção da variação no tempo, os resultados da variável faixa etária indicam um processo de variação estável, sem perspectiva de mudança.

[18] Em minha dissertação de Mestrado (CAVALCANTE, 2007), além da análise quantitativa mais detalhada, forneço uma proposta teórica de análise das estruturas negativas [Não VP não] e [VP não] de acordo com o modelo gerativo.

11 A concordância de número

Alan Baxter

O tema da concordância de número plural (doravante PL) no SN[1] é familiar no âmbito das discussões sobre as origens do Português Brasileiro (PB), já desde o século XIX. Na atualidade, há duas linhas de pensamento sobre as fontes dessa variação no PB: a linha *contatista* e a linha *derivista*.

A linha contatista atribui um papel fundamental aos contatos linguísticos históricos no Brasil e à atuação de potenciais processos de crioulização no desenvolvimento do PL no SN, e correlaciona essa variável com outros fenômenos morfossintáticos no PB considerados indícios de reestruturações que teriam as mesmas fontes (GUY, 1981a, 1981b, 2005; HOLM, 1987, 1992; BAXTER, 1998; BAXTER; LUCCHESI, 1999; LUCCHESI, 2000a). Para Guy (1981a), a marcação de PL é principalmente um fenômeno de estrutura linear: a primeira posição no SN é fundamental para a marcação de PL, mas as posições progressivamente para a direita são desfavoráveis à marcação;[2] havendo uma correlação entre classe gramatical e posição linear.

Em contrapartida, a linha derivista aponta para a presença de um padrão de concordância variável no SN em variedades do Português Europeu (PE) que teriam fornecido modelos para o PB (NARO; SCHERRE, 1993, 2000, 2007). As populações de falantes de português L2 teriam exacerbado esse modelo europeu. Aliás, contestando as afirmações de Guy (1981b), Scherre (1988, 1998), com base em dados da fala do Rio de Janeiro, argumenta que posição linear e classe gramatical não explicam o perfil da variação operante no PB (veja-se seção 11.3, abaixo).

Neste capítulo, retomamos alguns aspectos centrais desse debate e perguntamos:

---

[1] Neste trabalho, a menos que seja explicitado o contrário, o termo tradicional SN (sintagma nominal) será tratado como sinônimo do termo atual SD (sintagma determinante) que é definido na seção 11.2.

[2] Segundo Guy (1981a, p.168), a distribuição no SN dos itens com PL em dados do português da classe operária (corpus MOBRAL) do Rio de Janeiro registrou uma forte presença na posição inicial (posição 1 = 95% marcados), mas, a partir da segunda posição, houve um forte declínio na marcação: posição 2 (28%), posição 3 (22%), posição 4 (13%) e posição 5 (0%).

(i) até que ponto um modelo baseado na fala urbana dá conta da variação PL em variedades de português faladas por afrodescendentes em contextos rurais?

(ii) se essas variedades não poderiam ter desenvolvido a variação PL, havendo ou não variação nos modelos europeus?

Iremos investigar o perfil da variável PL em duas comunidades rurais de afrodescendentes, uma brasileira: a comunidade de Helvécia (Bahia), e uma africana: a comunidade dos tongas,[3] da roça Monte Café, na República de São Tomé e Príncipe.

11.1 Enquadramento metodológico

A comparação desses dois dialetos é motivada por uma série de fatos linguísticos e sociolinguísticos compartilhados:

(i) manifestam a variação na concordância de número,[4] como nos seguintes exemplos:

(1) a. *as pessoa de la é muito bom.* [Helvécia-M3]

b. *bota duas culeres n'agua morna* [Helvécia-M2]

(2) a. *os ôtros disse* [Monte Café-H3]

b. *issos coesa tudo* [Monte Café-H3]

(ii) são variedades de português que surgiram em microssociedades de trabalho 'controladas' e bastante fechadas: escravatura, num caso, e trabalho contratado, no outro;

(iii) surgiram em contextos demográficos com altas proporções de africanos e afrodescendentes contratados para administradores (BAXTER; LUCCHESI, 1999; BAXTER, 2002);

(iv) os seus antepassados africanos adquiriram o português pelo contato, principalmente a partir de modelos falados por colegas trabalhadores e, em parte, a partir de modelos falados pelos administradores;[5]

---

[3] Os tongas são os descendentes de africanos contratados, no final do século XIX e durante o século XX, para trabalhar nas grandes roças de café e cacau por períodos de 3 a 5 anos, com contratos renováveis. Os filhos nascidos em São Tomé passaram a ser a propriedade *de facto* das grandes roças. Para diversos aspectos da sócio-história dos tongas, vejam-se Rougé (1992) e Baxter (2002, 2004).

[4] Além de outras variáveis morfossintáticas, incluindo a variação nas concordâncias de Sujeito-Verbo e SN gênero (Baxter 2002, 2003).

[5] Também é possível que o português de São Tomé, uma variedade influenciada pelo crioulo de São Tomé, tivesse exercido alguma influência sobre o português dos tongas em contatos fora da roça.

(v) houve uma presença de português falado como segunda língua (L2) em fases anteriores;[6]

(vi) houve presença de línguas africanas dos grupos kwa e banto.[7]

Na discussão seguinte, recorreremos a duas fontes de análises: o estudo de Andrade (2003) do SN PL no dialeto de Helvécia e uma análise nossa da mesma comunidade, que incorpora dados de informantes adicionais gravados durante a mesma recolha que forneceu o *corpus* estudado por Andrade. A nossa análise reconfigura algumas das variáveis independentes contempladas por Andrade, para facilitar comparações com os dados dos tongas. Andrade analisou dados de 18 informantes, equitativamente distribuídos em três faixas etárias e por gênero: faixa 1: de 21 a 40 anos; faixa 2: de 41 a 60 anos; faixa 3: de mais de 60 anos. Em contrapartida, a nossa reanálise contemplou quatro faixas etárias: faixa 1: de 21 a 40 anos; faixa 2: de 41 a 60 anos; faixa 3: de 61 a 80 anos; e faixa 4: de mais de 85 anos. Nesse caso, as faixa 1 a 3 continham 6 informantes cada uma, e a faixa 4 só tinha 3 informantes. O *corpus* dos tongas de Monte Café tem a mesma estrutura.

Os dados foram submetidos ao programa VARBRUL na versão GOLDVARB-X (SANKOFF; TAGLIAMONTE; SMITH, 2005) para avaliar o efeito de um conjunto de variáveis independentes, linguísticas e sociais, sobre a variável dependente *morfema plural* (doravante morfema PL). Nessas análises, foram contempladas as variáveis independentes: (i) saliência fônica do morfema PL; (ii) marcas de plural precedentes ao item em análise; (iii) configuração estrutural do sintagma nominal; (iv) gênero do falante; (v) faixa etária do falante. O propósito do presente capítulo é comentar principalmente os resultados da variável linguística *posição linear do item PL em relação ao núcleo e função*, que foi a primeira variável independente selecionada pelo programa VARBRUL nas análises dos dois dialetos em questão.[8]

[6] A presença de falantes de português L2 nas primeiras décadas da Colônia Leopoldina (precursora de Helvécia) é documentada em Lucchesi e Baxter (2006). Por outra parte, Baxter (2002, 2004) registra a forte presença de falantes de português L2 na roça Monte Café até a década de 1940.

[7] Na Colônia Leopoldina, esses grupos poderiam ser mais equilibrados em determinadas roças (BAXTER; LUCCHESI, 1999), enquanto no caso dos tongas prevaleceu o grupo banto (BAXTER, 2004). Além das línguas africanas, houve também duas possíveis fontes de influência crioula, de menor importância: o crioulo de São Tomé, talvez presente na roça no século XIX entre ex-escravos crioulos; e o crioulo de Cabo Verde, presente minimamente no século XX. Contudo, há poucas evidências de aquisição dessas línguas por parte dos tongas. No caso da Colônia Leopoldina, na documentação histórica incompleta, registra-se também a presença de pelo menos um escravo de São Tomé, embora não se possa verificar se era são-tomense ou simplesmente "embarcado" em São Tomé. Na cidade de Salvador novecentista, os escravos dessa origem constituíam uma pequena minoria (SOUZA ANDRADE, 1988, p.98).

[8] Devido a limitações de espaço, a variável *saliência fônica do morfema de plural* não será comentada neste capítulo. Para estudos dessa variável nos dois dialetos em causa, remetemos o leitor para Andrade (2003) e Baxter (2004).

Em seguida, conjugamos essa discussão com o comentário da variável *marcas precedentes de PL* — considerada por Scherre (1989, 1998) como essencial para a caraterização da variável PL — e de alguns aspectos das variáveis sociais, sobretudo *faixa etária, 'etnia' dos pais,* e *gênero*.

11.2 Enquadramento teórico

A discussão da configuração estrutural da concordância e a inserção do morfema PL neste capítulo baseia-se na noção da estrutura do SD (ABNEY, 1987) que define o SN (material lexical) como complemento da classe D (determinante), e apoia-se nas perspectivas de Picallo (1991) e Cinque (1994), que postulam uma categoria funcional SNúm entre D e NP. A concordância (de número, neste caso) é considerada um reflexo de uma relação sintática local (CHOMSKY, 2001b), computada entre os elementos do SD.

Além dessas noções, serão relevantes elementos da perspectiva de Costa e Figueiredo Silva (2006) sobre a morfologia distributiva e não distributiva, e o Modelo 4-M de Myers-Scotton e Jake (2000a, 2000b). Com base nos trabalhos de Galves (1993), Scherre (1994) e Embick e Noyer (2001), Costa e Figueiredo Silva (2006) propõem que o PE e o PB manifestam diferentes realizações do morfema associado com a pluralidade. Assim, os morfemas podem se adjungir a núcleos independentemente do processo de movimento, e os morfemas podem ser realizados de duas maneiras: como morfemas *dissociados*, que não figuram na sintaxe propriamente dita, mas são inseridos após *spell-out*, como um reflexo de configurações sintáticas; ou como morfemas *singleton*, que se adjungem aos elementos funcionais que ancoram a informação sobre o número e a referência. Por outra parte, a hipótese de Myers-Scotton e Jake (2000a, 2000b) não gerativista, visualiza diferentes tipos de morfemas ativados em diferentes fases da produção linguística, alguns mais centrais ao sistema gramatical do que outros.

Por outro lado, na discussão dos motivos que estariam na base da variação PL nesses dois dialetos do português, lançaremos mão das perspetivas teóricas atuais nos estudos de contatos entre línguas (*language contact*) e aquisição de L2 (doravante SLA). Como as duas comunidades têm em comum a presença de português L2 nas suas histórias, cabe perguntar qual seria o contributo dos processos de aquisição de segunda língua para a constituição da variável em estudo. Nesse sentido, a discussão das análises é orientada pelas teorias sociolinguísticas que captam as situações de contato que envolvem mudança de língua via uma fase intermediária de SLA, com fases de bilinguismo, e a aquisição de L1 com *input* de

modelos de L2.[9] Também, é importante notar que esse processo de aquisição envolve uma mudança de alvo. Em fases iniciais em contingentes *escravizados*, a L2 do escravizado pode ter diversas fontes: (i) a L1 dos administradores, (ii) a L2[10] de outros escravos, baseada na L1 dos administradores, e, eventualmente, (iii) uma L1 da população escravizada baseada na L2 dos escravos (Smith, 2006). Além disso, a depender das circunstâncias sociodemográficas, é importante sublinhar que (i) a L2 do escravo pode constituir um alvo significativo de aquisição em determinados períodos, e (ii) a nova variedade de L1 da comunidade *assimilante* pode tornar-se o alvo primário de aquisição por parte de crianças. Desde que sejam mantidas as condições sociais restritivas (a escravidão, o trabalho contratado, o isolamento), e o forte desequilíbrio sociodemográfico de proporção de escravo para não escravo, a variedade de L1 exterior à comunidade assimilante, falada por administradores e livres, constitui um alvo periférico.[11] Circunstâncias semelhantes são observadas na atualidade em comunidades minoritárias de trabalhadores imigrantes estrangeiros na Europa (CORNIPS; HULK, 2006).[12]

Em situações de mudança de língua,[13] material de L2 pode ser incorporado na nova L1 da comunidade nas primeiras fases da sua constituição e transmitido por gerações posteriores (WINFORD, 2003, p.245-247; CORNIPS; HULK, 2006; MONTRUL, 2006). Ao mesmo tempo, a presença de bilinguismo pode também facilitar a transferência de elementos das L1 originais para a segunda (e nova) L1 da comunidade (BAO, 2005), assim como convergências funcionais entre as duas línguas (MONTRUL, 2006).

Entre os fatores que influenciam a forma da L2 de aprendizes adultos, os seguintes são centrais nas investigações atuais:

---

[9] São situações captadas pela noção de *transmissão irregular*. Em realidade, essa noção vem ao encontro do conceito de *group second language acquisition or shift* "aquisição de L2 por um grupo, ou mudança de língua" (WINFORD, 2003, p.235-255), que são essencialmente conceitos paralelos.

[10] Por L2 entende-se 'variedades de interlíngua'. Haveria um forte elemento de variação individual, mas simultaneamente haveria processos de nivelamento das divergências individuais, induzidos pela comunicação.

[11] São muito pertinentes as críticas de Smith (2006) à suposição tradicional de que o escravo numa situação dessas desejaria mesmo aprender a língua do administrador. Na opinião desse linguista, e considerando as investigações sobre os aprendizes de L2 em situação de contato, é mais provável o escravo simplesmente contentar-se em aprender o mínimo necessário para poder comunicar-se com os administradores, sendo o mais importante poder comunicar-se com outros escravos.

[12] Estes investigadores estudam a perda do gênero no artigo definido singular no holandês falado por descendentes de trabalhadores imigrantes. A reestruturação do artigo tem as suas bases no holandês L2 dessas comunidades minoritárias.

[13] Quando uma comunidade muda de língua, abandonando uma língua A e adquirindo uma língua B.

1. **Simplificação**: presença de formas reduzidas e regularizadas (KLEIN; PERDUE, 1997; MATHER, 2000, 2006; SIEGEL, 2004), em função da natureza do *input* (p. ex., via frequências, saliências; transparência semântica e pragmática) e da Gramática Universal (GU);
2. **Transferência e relexificação**:[14] formas estruturais da L1 do aprendiz transferidas para a L2 incipiente (SIEGEL, 2006; SCHWARTZ; SPROUSE, 1994; SPROUSE, 2006);
3. **SLA imperfeita**: por exemplo a *Failed Functional Features Hypothesis* (FFFH) (HAWKINS; CHAN, 1997; FRANCESCHINA, 2002, 2003), segundo a qual o aprendiz adulto seria incapaz de adquirir traços das categorias funcionais parametrizadas, uma incapacidade condicionada pelo período crítico.[15]

Em essência, embora a GU e questões processuais independentes da L1 do aprendiz estejam envolvidas nos fatores 1 e 3, também pode estar implicada neles a influência da L1: na saliência e na transparência semântica, no fator 1; e na impossibilidade de adquirir categorias funcionais novas/diferentes, no fator 3. Ao mesmo tempo, consideramos pertinentes os fatos aportados pelos recentes estudos da aquisição L1 da morfologia de PL no PB, que apontam para o papel central das classes morfológicas fechadas (*closed class morphemes*) pré-nominais na orientação do desenvolvimento do sistema de PL (CORRÊA; AUGUSTO; FERRARI-NETO, 2005).

Os referidos fatores aquisicionais são especialmente relevantes para a forma e o processamento dos dados linguísticos primários (DLP) disponíveis para a aquisição de L1 entre as gerações mais antigas de Monte Café e da Colônia Leopoldina.

11.3 Posição e classe gramatical

Diversos estudos sobre o PB têm apresentado a posição linear do item pluralizável como chave para a flexão de PL. Guy (1981b) aponta para uma conexão entre posição linear do item e a sua classe gramatical,[16] e localiza o eixo da marcação

[14] Segundo Siegel (2006, p.31), a *relexificação* copia na L2 a entrada lexical da L1, concedendo-lhe uma forma nova, e mantém todas as propriedades da estrutura da L1; já a *transferência* envolve o uso de morfemas da L2 com somente algumas das propriedades dos morfemas correspondentes da L1.

[15] Em essência, a FFFH também aponta para a influência da L1.

[16] Haveria um paralelismo de efeito entre a influência dos determinantes e a primeira posição, dos substantivos e a segunda posição, dos adjetivos e a terceira posição (GUY, 1981b, p.90).

de PL nos determinantes, configuração atribuída ao substrato kwa e banto (GUY, 1981a, 1981b, 2005). Contestando as afirmações de Guy, Scherre (1988, p.152-153; 1998, p.94-97) afirma que a análise distribucional das classes em dados urbanos do Rio de Janeiro não revela aquela correlação, porque há uma diversidade de classes gramaticais em cada posição no SN. Fundamentalmente, essa autora conclui que:

- a primeira posição favorece a presença de marca de plural em qualquer classe gramatical;
- os possessivos e demais elementos não nucleares podem também ocorrer na segunda posição, com frequências mais elevadas do que os substantivos;
- os adjetivos também são muito frequentes na segunda posição, e se apresentam mais marcados na segunda posição do que na terceira;
- os substantivos também ocupam a terceira posição com uma frequência maior do que a dos adjetivos.

Apesar da validade dessas afirmações, o quadro da análise distribucional (SCHERRE, 1988, p.154) ainda sugere alguma relação entre posição linear e classe gramatical.[17] Os itens em posição pré-nuclear (= pré-nominal) registram frequências mais altas do que o substantivo em terceira posição (p.154). Portanto, a noção de um paralelismo entre posição e classe ainda merece consideração, embora as posições 1, 2 e 3 não estejam ocupadas exclusivamente por determinantes, substantivos e adjetivos, respetivamente.

Contudo, a polêmica em torno do efeito da posição e a classe se desfaz quando se considera que a linearidade encobre SNs com estruturas internas diferentes, por exemplo: Det+N+Adj; Det+Det+N; N+Adj. No fundo, essa visão é inerente à reanálise adotada por Scherre (1988, 1998), quando demonstra que as relações entre os itens do SN oferecem uma melhor explicação da variação do que somente a linearidade. Na reanálise (1988), contempla como variável complexa o cruzamento de posição nuclear, classe nuclear e não nuclear, e a relação linear entre classe nuclear e não nuclear. Os resultados gerais (SCHERRE, 1998, p.100-

[17] Nos dados de Scherre (1988, p.154), os artigos e demonstrativos constituem a esmagadora maioria dos itens na posição 1, constituindo 3.268 do total dos 4.230 itens marcados para PL nessa posição. Os outros itens marcados nessa posição apresentam números brutos abaixo de 450, e geralmente abaixo de 200. Na segunda posição, a esmagadora maioria dos itens marcados, são os substantivos, que constituem 2.777 do total de 5.778 itens marcados, e, nessa posição também, o adjetivo, embora marcado com uma frequência de 77%, só registra 125 itens marcados. Aliás, embora em frequência o substantivo em segunda posição seja pouco marcado em comparação com a maioria das categorias que aparecem nessa posição com frequências acima de 65%, é evidente que a maioria dessas outras categorias (essencialmente itens pré-nucleares) só entram em jogo quando o núcleo está em terceira posição. Na posição três, o substantivo com marcação de PL é o item marcado mais numeroso (317 do total de 968 itens marcados), e o adjetivo é o segundo item marcado mais numeroso (125 do total de 968 itens marcados).

103) indicam que (i) qualquer classe anteposta ao núcleo do SN favorece a marcação de PL, enquanto qualquer classe posposta ao núcleo a desfavorece; (ii) o substantivo só favorece a marcação quando está em primeira posição.[18] Ao mesmo tempo, Scherre observa que a presença ou ausência do morfema PL em um item do SN favorece a sua presença ou ausência, respectivamente, no item seguinte. Nessa perspectiva, para modelar a variação PL, é imprescindível contemplar as duas variáveis: *configuração do SN* e *presença de marca precedente*.

11.4 Configuração do SN e marcação do plural em dois dialetos afro-lusófonos

O estudo de Andrade (2003) da variação PL no dialeto de Helvécia retoma vários aspectos da proposta de Guy (1981a, 1981b). Em primeiro lugar, Andrade (p.95-96) confirma, na análise distribucional do PL, a correlação entre classe gramatical e posição linear, porque esse dialeto apresenta um SN com uma estrutura muito mais simples e regular do que aquela dos dados urbanos analisados por Scherre (1988) e Guy (1981a): a maioria dos SN PL apresenta só dois elementos: Det + N. Na primeira posição do SN, quase 77% dos itens são determinantes, com uma frequência de marcação de PL superior a 94% (artigos, pronomes possessivos, demonstrativos). Na segunda posição, 91% dos itens são substantivos e 0,5%, adjetivos, mas só 8% dos substantivos estão pluralizados, contra 25% dos adjetivos (ANDRADE, 2003, p.93). Nesse dialeto, 66% de todos os adjetivos aparecem na terceira posição, com apenas 3% de marcação de PL. Em segundo lugar, a análise global da variável complexa proposta por Scherre, *posição em relação ao núcleo do SN*, revela quatro tendências centrais: (i) forte marcação no pré-núcleo, (ii) forte marcação do núcleo em primeira posição, (iii) forte ausência de marcação no núcleo com determinantes, e (iv) forte ausência de marcação no pós-núcleo.

Os resultados obtidos por Andrade (2003) nos animaram a efetuar uma análise comparada, que estudou o perfil de quatro faixas etárias nos dialetos de Helvécia e Monte Café. A variável *posição do item em relação ao núcleo do SN* foi configurada seguindo as linhas gerais do modelo de Scherre (1988, 1998), porém distinguimos as posições pré-nucleares em termos de adjacência ao núcleo, um procedimento adotado por Lopes (2001) na sua análise da fala de Salvador. Os dados foram processados pelo pacote estatístico GOLDVARB X (SANKOFF; TAGLIAMONTE; SMITH, 2005). Para os dois dialetos, foram feitas análises

[18] Contudo, nesta posição o núcleo é menos marcado do que as classes antepostas que ocupam a primeira e segunda posição dentro do SN. Aliás, o núcleo em terceira posição é menos desfavorável à marcação de PL (.39) do que em segunda posição (.24) (SCHERRE, 1998).

independentes de cada faixa etária e uma análise global de todas as faixas em conjunto (Tabela 1).[19] Essa variável estrutural foi selecionada como a mais relevante para o condicionamento do uso do PL.

As análises independentes de cada faixa etária são especialmente instrutivas para a compreensão do desenvolvimento da variável PL e o seu enquadramento linguístico nas duas comunidades. Sobretudo, demonstram que as diversas faixas etárias têm gramáticas diferentes relativamente à estrutura do SN e à atribuição de marcas de PL. Aliás, determinadas tendências das faixas avançadas são preservadas nas faixas posteriores. Nos dois dialetos, os perfis da distribuição por faixas etárias apontam para processos de mudança aquisicional, uma perspectiva que será confirmada na análise das variáveis sociais na seção 11.6, mais adiante.

Comparando os resultados da análise do dialeto de Helvécia (Tabela 1, seção A) com os da análise do dialeto dos tongas (Tabela 1, seção B), é evidente que há diversos paralelos nos papéis da estrutura pré-nuclear, nuclear e pós-nuclear na marcação de PL, e ainda no desenvolvimento do sistema de marcação de PL de faixa para faixa nos dois dialetos:

(i) Em todas as faixas etárias, há um declínio no grau de marcação da esquerda para a direita do SN; há muito mais marcação de PL na área pré-nuclear do que no núcleo ou nas posiçoes pós-nucleares. Esse efeito é mais aparente nas faixas etárias avançadas. A posição pré-nuclear adjacente ao núcleo desempenha um papel central, sobretudo na faixa 4 de Helvécia e nas faixas 3 e 4[20] dos tongas que apontam para um sistema cujo mecanismo predominante de marcação de PL está radicado na estrutura funcional (artigo, demonstrativo, possessivo, quantificador). Aliás, nessas mesmas faixas etárias, há predominantemente um único elemento nessa posição.

Na faixa 4 dos tongas, dos 15/18 dados pré-nucleares com marcação de PL, nove artigos, dois demonstrativos e um quantificador levam PL em sintagmas cujos núcleos nominais não levam PL. Na faixa 4 de Helvécia, há 129 dados pré-nucleares adjacentes, todos itens funcionais em primeira posição do sintagma, e 125 levam PL. Desses, 71/72 artigos levam PL, enquanto todos os 42 demonstrativos, os 9 possessivos, um único quantificador e 2/3 indefinidos levam PL, todos eles em sintagmas cujos núcleos nominais só levam PL em 3/158 casos.

---

[19] Como o objetivo do presente estudo foi discutir a posição do item marcado em relação ao núcleo, não nos debruçamos sobre a aplicação total da regra de concordância de número. Contudo, esse dado pode ser de interesse para o leitor: Helvécia revela uma aplicação da regra em apenas 9% dos casos, enquanto o dialeto dos tongas revela um maior uso da concordância total no valor de 35,6%.

[20] Não foi possível realizar uma análise probabilística dos dados da F4 dos tongas, por ser categórica a ausência de variação (ausência de PL) na maioria dos fatores das variáveis independentes. Para efeito da discussão estão incluídos os dados do arquivo CELL.

Tabela 1 - Efeito de <posição e classe gramatical> sobre o PL no SN - Faixas etárias e análise de conjunto comparadas

(A) DIALETO DE HELVÉCIA	Faixa 1 (20 a 40 anos) *Significance = .010* *Log likelihood =-250.299*			Faixa 2 (41 a 60 anos) *Significance=.000* *Log likelihood =-175.023*			Faixa 3 (61 a 80 anos) *Significance = .007** *Log likelihood =-64.037*			Faixa *4* (85 a 103 anos) *Significance = .003* *Log likelihood =-54.849*			Análise global (as quatro faixas) *Significance = 0.048* *Log likelihood =-538.840*		
FATOR	Nº PL	% PL	P.R.	Nº PL	% PL	P.R.	Nº PL	% PL	P.R.	Nº PL	% PL	P.R .	Nº PL	% PL	P.R .
(1) 1ª posição à esquerda, adjacente ao núcleo	357/365	97.8%	**0.919**	386/394	98%	**0.939**	159/163	97.5%	**0.960**	125/129	96.9%	**0.991**	1027/1051	97.7%	**0.914**
(2) 2ª posição à esquerda do núcleo, não adjacente	39/43	90.7%	**0.673-**	25/32	78.1%	**0.421**	8/8	100%	--	2/4	50%	--	74/87	85.1%	**0.479**
(3) Núcleo em 1ª posição	0/1	0%	--	0/0	100%	--	0/1	0%	--	0	0%	--	0/1	0%	--
(4) Núcleo em 2ª posição	127/550	23.1%	**0.153**	38/510	7.5%	**0.179**	9/241	3.7%	**0.111**	4/166	5%	**0.028**	121/1414	8.6%	**0.153**
(5) Núcleo em 3ª posição em diante	9/46	32.7%	**0.174**	2/42	4.8%	**0.001**	5/16	31.23%	**0.268**	3/10	30%	--	19/114	16.7%	**0.397**
(6) Itens à direita do núcleo	0/23	0%	--	0/26	0%	--	0/21	0%	--	0/20	0%	--	0/90	0%	--

(B) DIALETO DOS TONGAS	Faixa 1 20 a 40 anos) *Significance = .024* *Log likelihood =-573.605*			Faixa 2 (41 a 60 anos) *Significance=.005* *Log likelihood =-418.641*			Faixa 3 (>61 anos) *Significance = .036* *Log likelihood =-128.592*			Faixa 4 (>80 anos)**			Análise global (as quatro faixas) *Significance = .000* *Log likelihood =-1156.674*		
FATOR	Nº PL	% PL	P.R.	Nº PL	% PL	P.R.	Nº PL	% PL	P.R .	Nº PL	% PL	P.R.	Nº PL	% PL	P.R.
(1) 1ª posição à esquerda, adjacente ao núcleo	533/543	98.2%	**0.835**	296/309	95.8%	**0.913**	158/181	87.3%	**0.958**	12/18	83.3%	--	993/1042	95.3%	**0.915**
(2) 2ª posição à esquerda do núcleo, não-adjacente	65/67	97%	**0.770**	26/27	96.3%	**0.884**	0/2	0%	--				91/96	94.8%	**0.795**
(3) Núcleo em 1ª posição	33/35	94.3%	**0.735**	6/6	100%	--	1/2	50%	--				40/43	93.0%	**0.798**
(4) Núcleo em 2ª posição	402/659	61%	**0.237**	158/597	26.5%	**0.243**	22/487	4.5%	**0.260**	0/71	0%	--	582/1805	32.2%	**0.223**
(5) Núcleo em 3a posição em diante	61/96	63.5%	**0.273**	20/53	37.7%	**0.320**	0/13	0	--				81/162	50%	**0.267**
(6) Itens à direita do núcleo	30/76	39.5%	**0.144**	3/27	11.1%	**0.084**	1/15	6.7%	**0.024**	0/2	0%	--	34/120	28.3%	**0.112**

** Não foi possível realizar uma análise VARBRUL da faixa 4. Contudo, uma rodada realizada com a inserção de um dado falso de marcação de PL no fator (4) revela um P.R. de 0.999 para o fator (1) <1ª posição à esquerda, adjacente ao núcleo>.

Tanto nos tongas como nos dados de Helvécia, nota-se que o PL é atribuído predominantemente em itens funcionais que indicam a referencialidade do SN.

Nas outras faixas etárias, conforme diminui a idade, percebe-se um acréscimo na estrutura pré-nuclear, com a adição do item flexional não adjacente ao núcleo. Na transição da faixa 2 para a faixa 1 dos dois dialetos, essa posição se torna relativamente significativa (Helvécia, faixa 2: p.r. 0.421, faixa 1: p.r. 0.673; tongas, faixa 2: p.r. 0.884, faixa 1: p.r. 0.770). O forte papel da posição adjacente ao núcleo mantém-se ao longo das sucessivas faixas etárias e torna-se uma caraterística dessas comunidades (cf. análises globais, Tabela 1 seções (A) e (B), FATOR (1)).

(ii) O núcleo em segunda posição é bastante desfavorável à marcação de PL. Nas faixas 4, notam-se a categórica ausência de PL nos dados dos tongas e a sua quase categórica ausência nos de Helvécia (3/158, p.r. 0.028). Em faixas etárias sucessivas, esse perfil da segunda posição se mantém e continua sendo bastante desfavorável à marcação de PL nos dois dialetos: nos dados de Helvécia, nota-se um leve decréscimo nesse desfavorecimento (faixa 4: p.r. 0.028, faixa 3: p.r. 0.111, faixa 2: p.r. 0.179, faixa 1: p.r. 0.153) enquanto no dialeto dos tongas o peso se mantém com um valor entre 0.2 e 0.3 (faixa 3: p.r. 0.260, faixa 2: p.r. 0.243, faixa 1: p.r. 0.237). Nos dois dialetos, o núcleo na terceira posição em diante é também bastante desfavorável à marcação de PL.

(iii) Os itens pós-nucleares desfavorecem muito a marcação de PL. No dialeto de Helvécia, nas quatro faixas,[21] há uma categórica ausência de marcas de PL nas posições pós-nucleares. No dialeto dos tongas, essa posição é altamente desfavorável à marcação, mas manifesta um enfraquecimento gradual, conforme diminui a idade (faixa 3: p.r. 0.024, faixa 2: p.r. 0.084, faixa 1: p.r. 0.144).

(iv) Há uma série de diferenças gramaticais no SN das sucessivas faixas etárias. Detecta-se uma complexificação estrutural. As faixas avançadas (faixa 4, Helvécia e tongas) utilizam predominantemente estruturas com duas posições-funções: um único elemento pré-nuclear. A faixa 4 dos tongas manifesta um SN mínimo, de duas categorias: DET + N. As estruturas com três itens — uma pequena minoria nas faixas 3 dos dois dialetos e na FAIXA 4 de Helvécia — são principalmente do tipo DET + X + N ou DET + N + X, prevelecendo esta última. As faixas mais jovens têm uma estrutura pré-nuclear mais complexa e, por isso, o núcleo e o pós-núcleo ocupam mais posições. Também, entre os tongas, a faixa 1 desenvolve a marcação do núcleo em primeira posição absoluta do SN, elemento incipiente nas faixas 3 e 2. Por outra parte, no dialeto de Helvécia, a marcação do núcleo plural em primeira posição absoluta é nula no *corpus* considerado aqui.

[21] Andrade (2003, p.92) registra 2% de marcas de PL nessa posição, com um peso relativo de 0.00.

(v) As regras de atribuição de PL são diferentes nas sucessivas faixas etárias. Nas faixas mais avançadas, a marcação de PL é estabelecida primeiro no pré-núcleo, na posição adjacente ao núcleo (Helvécia, faixa 4: p.r. 0.991, faixa 3: p.r. .960; tongas, faixa 4: 83.3%, faixa 3: p.r. 0.958). Mas, nas faixas intermédias e jovens, onde há múltiplos itens pré-nucleares (veja-se <segunda posição à esquerda do núcleo, não adjacente> (Helvécia, faixa 2 para faixa 1; tongas, faixa 3, para faixa 2), a marcação de PL favorecida na posição adjacente se estende para a esquerda dessa posição. Aliás, essas mesmas faixas etárias manifestam uma expansão da marca para o núcleo em outras posições, mas a marcação dessas posições é ainda relativamente inibida.

Desses comentários, e apesar das diferenças entre os dois dialetos, é evidente que os paralelos no efeito *posição em relação ao núcleo* são fortes.

11.5 O efeito da presença de marcas precedentes sobre a marcação do plural

Um outro procedimento tradicional nos estudos sobre a variação PL no SN no PB é a análise do efeito da presença de outras marcas de PL no SN, anteriores ao item em análise (SCHERRE, 1988, 1989, 2001). Esses estudos encontram uma maior probabilidade de marcação de PL nos itens precedidos de outros itens com marca de PL. Por outro lado, a ausência do PL antecedente induz a ausência da marca no item analisado. Scherre atribuiu esse resultado a uma tendência funcionalista de *processamento paralelo*, resumida no princípio: *marcas levam a marcas e zeros levam a zeros*. Embora esse efeito tenha sido encontrado em outros estudos do PB, principalmente com dados urbanos, à luz dos resultados considerados na seção anterior, cabe perguntar se a variável *marcas precedentes* manifesta o mesmo efeito nos dialetos de afrodescendentes focalizados aqui. A Tabela 2 compara as análises relevantes.

No dialeto de Helvécia, a marcação de PL em segunda posição do SN manifesta uma semelhança parcial com os padrões observados por Scherre (1998, p.108) e Lopes (2001, p.211). Nesse sentido, os fatores que favorecem a marcação na segunda posição do SN são a ausência da marca formal na primeira posição ou a presença de um numeral nessa posição. Diferentemente dos resultados dos estudos de Scherre (1988, 1998), mas de modo semelhante aos de Guy (1981a) e Lopes (2001), a presença da marca formal na primeira posição não favorece a marcação do segundo item. Dessa maneira, a mínima marcação no item que é precedido por um item formalmente marcado, coincide com a proposta de Guy (1981a) de que o português

popular do Brasil tende a marcar o plural na primeira oportunidade (Andrade, 2003, p. 94). Por outro lado, a marcação do item em terceira posição é desfavorecida, tanto pela presença de uma mistura de marcas com marca formal precedente (p.r. .412), como pela presença da mistura de marcas com zero formal precedente (p.r. .099). Contudo, é evidente que o menor desfavorecimento de PL no primeiro desses dois fatores e o maior desfavorecimento no segundo poderiam ser interpretados como indícios de um leve efeito de processamento paralelo.

Nos resultados da análise dos dados dos tongas há paralelos com o dialeto de Helvécia (Tabela 2). Em segunda posição, o PL é favorecido pela ausência da marca formal na primeira posição (Helvécia: p.r. 0.844; tongas: p.r. 0.651). Também, a inibição da marca de PL em itens da terceira posição pela mistura de marcas com zero precedente apresenta um perfil paralelo (Helvécia: p.r. 0.099; tongas: p.r. 0.190), porém, o dialeto dos tongas manifesta três diferenças notáveis: (i) a marca precedente em primeira posição é favorável à marcação (p.r. 0.583); (ii) o numeral em primeira posição a desfavorece (p.r. 0.413); e (iii) a marcação do item em terceira posição é favorecida pela presença de uma mistura de marcas com marca formal precedente (Helvécia: p.r. 0.412; tongas: p.r. 0.557).

Tabela 2 - Influência do plural morfológico ou semântico sobre a marcação de PL no item subsequente – dialetos de Helvécia e Monte Café (análise global)

MARCAS PRECEDENTES	Helvécia Análise Global (quatro faixas etárias) *sig.* 0.048			Tongas (Monte Café) Análise Global (quatro faixas etárias) *sig.* 0.000		
	Número de ocorrências	%	p.r.	Número de ocorrências	%	p.r.
Marca formal na 1ª posição; item analisado na 2ª posição (*dos* ***meus*** *pai*)	77/1031	7.5	.499	481/1001	48.1	.583
Numeral na 1ª posição; item analisado na 2ª posição (*dez filhos*)	58/425	13.6	.641	194/936	20.7	.413
Ausência de marca formal na 1ª posição; item analisado na 2ª posição (*do* ***meus*** *pai*)	10/18	55.6	.844	16/45	35.6	.651
SN de três itens: mistura de marcas precedentes com marca formal; item analisado na 3ª posição (*seus quatro* ***meses***)	8/65	12.3	.412	68/130	52.3	.557
SN de três itens: mistura de marcas precedentes com zero; item analisado na 3ª posição (*toda essas* ***coesa***)	3/113	2.7	.099	3/44	6.8	.190

O que se depreende desses resultados é que o princípio de *marcas levam a marcas e zeros levam a zeros* não é relevante para o dialeto de Helvécia quando se trata de marcas flexionais precedentes. No dialeto dos tongas, o princípio se revela contraditório. Nesse caso, a marca flexional precedente na primeira posição leva à marca na segunda posição, e mistura de marcas precedentes com marca formal levam à marca na terceira posição. Porém, zero em primeira posição favorece a marcação da segunda posição, enquanto a mistura de marcas precedentes com zero favorece zero na terceira posição. O princípio parece marginal à atribuição de PL nesses dois dialetos que manifestam estruturas de SN e padrões de marcação de PL bastante simples. Aliás, atesta-se o fato de que diferentes dialetos podem desenvolver diferentes padrões de marcação no que diz respeito à influência das *marcas precedentes*, embora com semelhanças parciais.

Acreditamos que as diferenças entre os dois dialetos se devem aos substratos e aos contextos sociolinguísticos em que se desenvolveram. No caso do dialeto de Helvécia, houve uma presença mais equilibrada de falantes de línguas kwa e banto, e o iorubá só tem vestígios da tipologia concordante da família níger-congo. Por outra parte, entre os tongas poderíamos esperar mais sensibilidade para a noção da concordância graças à presença constante de línguas banto com sistemas de concordância morfologicamente ricos. Ao mesmo tempo, o papel privilegiado de alguns tongas, em trabalhos "especializados", teria facilitado os contatos com falantes de PE. Esses dois fatores poderiam favorecer o desenvolvimento da marcação de PL.

11.6 Variáveis sociais

Tendo visto algumas evidências de processos aquisicionais no perfil estrutural da variável PL nesses dialetos, cabe agora perguntar como a variável se enquadra nas duas comunidades de fala. No seu estudo de PL no dialeto de Helvécia, Andrade (2003) considera a influência de quatro variáveis extralinguísticas: *faixa etária*, *gênero*, *estada fora da comunidade* e *escolaridade*. Contudo, na discussão a seguir, nos concentramos primeiro sobre o efeito da faixa etária nas duas comunidades em consideração. Em seguida, contemplamos a influência que a origem dos pais dos falantes tongas exerce sobre a variável PL, refletindo brevemente sobre as implicações tanto para Monte Café como para Helvécia. Depois disso, por limitações de espaço, a discussão se encerra com alguns comentários resumidos sobre as variáveis *gênero* e *estada fora da comunidade* relativamente à comunidade de Helvécia, estabelecendo alguns pontos de comparação com a comunidade tonga da roça Monte Café.

Nas análises globais das duas comunidades, a *faixa etária* é a primeira variável social selecionada, embora precedida em importância sempre pela variável linguística

posição-função em relação ao núcleo. Nos resultados da Tabela 3,[22] o perfil da variável *faixa etária* sugere que a flexão de PL está em processo de mudança aquisicional nas duas comunidades, um fato já inferido na análise estrutural na seção 11.4.[23] Nas faixas mais avançadas, a marcação de PL é mínima. No caso do *corpus* de Helvécia, trata-se do tipo de falante que nasceu quando ainda havia na comunidade um certo número de ex-escravos, procedentes dos microcontextos das grandes roças da Colônia Leopoldina. São falantes que nasceram em pequenas roças em contextos semi-isolados quase exclusivamente de descendentes dos escravos, na quinta geração dessa comunidade depois da fundação da Colônia Leopoldina. Teriam como DLP, os modelos linguísticos afro-brasileiros da mesma comunidade.

Tabela 3 - Marcação de PL e faixa etária - Helvécia e Monte Café

FAIXAS ETÁRIAS	Helvécia Análise Global (quatro faixas etárias) *sig.* 0.048			FAIXAS ETÁRIAS	Monte Café Análise Global (três faixas etárias) *sig.* 0.000		
	Número de ocorrências	%	p.r.		Número de ocorrências	%	p.r.
Faixa 1 (21 a 40 anos)	475/973	48.8	0.638	Faixa 1 (21 a 40 anos)	1120/1469	76.2	0.792
Faixa 2 (41 a 60 anos)	451/1004	44.9	0.483	Faixa 2 (41 a 60 anos)	524/1118	46.9	0.425
Faixa 3 (60 a 80 anos)	181/269	40.2	0.393	Faixa 3 (61 a 80)	179/694	25.8	0.112
Faixa 4 (85 a 103 anos)	134/333	40.2	0.295	Faixa 4 (80+ anos)	12/85	14.1	0.095

Por outra parte, na comunidade tonga, a faixa 3 e também parte da faixa 2 correspondem a falantes nascidos de casais africanos ou casais mistos (= africano + tonga). Portanto, para o seu processo de aquisição de L1, essas pessoas teriam acesso a um forte componente de português L2 falado por africanos, um português aprendido em situação de contato, além do modelo de língua africana, nesse caso o umbundo. Evidências indiretas da relevância do português L2 de adultos são apresentadas em Baxter (2004, p.104-105), que avalia a marcação de PL no desempenho dos falantes das faixas 2 e 3, conforme a procedência dos pais: se nasceram na África ou na roça Monte Café. A Tabela 4 apresenta os resultados. Os falantes com pais nascidos na roça Monte Café favorecem a marcação de PL, p.r. .63, enquanto os falantes com pais africanos desfavorecem a marcação de PL, p.r. .36.

---

[22] Andrade (2003, p.124) analisa a faixa etária em relação à aplicação plena da regra de concordância e também o uso do morfema PL, obtendo resultados que apontam para o mesmo perfil etário observado na presente análise na Tabela 3.

[23] Esta perspetiva coincide com os resultados das análises de Andrade (2003, p.124) e Baxter (2004, p.104).

Tabela 4 - Efeito da origem dos pais dos falantes tongas (faixas 2 e 3), sobre a marcação de PL

ORIGEM DOS PAIS	NÚMERO DE OCORRÊNCIAS/TOTAL	%	PESO RELATIVO
África	275/914	30	0.36
Monte Café	440/983	45	0.63

sig. 0.029

Disso, deduz-se que a presença africana teve um efeito sobre a não marcação de PL, um efeito que poderia ter origens no português L2. Na sócio-história das duas comunidades, houve situações sociais relativamente confinadas, presença de falantes de línguas africanas e falantes de português como L2, aprendizagem de português L1 com base em modelos de português L2 e de português L1 cuja variedade predominante era falada pelos afrodescendentes das mesmas comunidades.

Por outra parte, o maior grau de marcação de PL nas faixas etárias jovens que se observa na Tabela 3 poderia estar relacionado com a influência de diversos fatores socioeconômicos atuantes na segunda metade do século XX que contribuíram para um nivelamento linguístico. No caso de São Tome, houve, a partir da década de 1940, uma ligeira liberalização do sistema de administração das roças, com mais liberdade de movimento para os serviçais; e os filhos de serviçais tiveram acesso à escolarização. Com isso, caraterísticas linguísticas das gerações anteriores estariam sendo eliminadas em função de um uso mais padrão, sob influências linguísticas de fora da comunidade. Já para a comunidade de fala de Helvécia, são significativos os seguintes fatores: a migração periódica dos homens à procura de trabalho, a abertura de rodovias, a presença das grandes empresas agroindustriais de celulose, a chegada dos meios de comunicação de massa e a educação pública (LUCCHESI, 2000a, p.286).

Para a verificação de uma possível mudança aquisicional no dialeto de Helvécia, Andrade (2003) recorre principalmente a análises das varáveis extralinguísticas *gênero* e *estada fora da comunidade*. Os resultados da variável *gênero* estão registrados na Tabela 5. Embora a diferença entre os *gêneros* pareça mínima em termos de frequência de uso, os pesos relativos esclarecem a realidade. Os homens favorecem (modestamente) tanto o uso da regra sintática da concordância como da morfologia flexional de número, e as mulheres desfavorecem as duas variáveis.

Tabela 5 - Helvécia: efeito do gênero do falante sobre a variável PL

GÊNERO	NÚMERO DE OCORRÊNCIAS/TOTAL	%	PESO RELATIVO
Feminino	642/1451	44	.43
Masculino	668/1443	46	.57

Fonte: Andrade, 2003, p.105.

Esse perfil encontra paralelos nos estudos de outras variáveis sociolinguísticas dessa comunidade. Os homens têm um maior grau de integração social, o que implica um maior contato com padrões linguísticos adventícios. Passam mais tempo fora da comunidade (deslocam-se para outras cidades da região para vender produtos das roças, ou para trabalhar por temporadas e também trabalham em empresas que funcionam na região, saindo diariamente da comunidade), enquanto as mulheres tendem a estar mais enraizadas na sua própria comunidade, dedicadas à roça e ao trabalho doméstico (ANDRADE, 2003; BAXTER; LUCCHESI; GUIMARÃES, 1997; LUCCHESI, 2000a). Andrade (2003, p.106) observa que a variável *estada fora da comunidade* registra resultados paralelos aos da Tabela 5, sublinhando mais ainda a relevância do contato e convívio fora da comunidade para o processo de aproximação do dialeto aos padrões de maior prestígio social.

Na comunidade tonga de Monte Café, encontramos uma situação semelhante relativamente à variável *gênero*, como se pode ver na Tabela 6.

Tabela 6 - Tongas: influência do gênero do falante sobre o uso do morfema PL

GÊNERO	NÚMERO DE OCORRÊNCIAS/TOTAL	%	PESO RELATIVO
Feminino	724/1379	52.5	.401
Masculino	1094/1883	58.1	.573

sig. 0.000

Os fatos sociais que estão por trás desse perfil são semelhantes àqueles de Helvécia. Os homens tongas, embora sempre residentes na comunidade de Monte Café, têm uma maior integração social e, portanto, um maior contato com padrões linguísticos adventícios. Esse é um processo que tem as suas raízes na época colonial, quando o homem tonga desempenhou um papel intermediário na comunicação entre os administradores portugueses e os serviçais africanos. Aliás, graças a esse estatuto de intermediário, o homem tonga podia ter contatos fora da comunidade. A mulher tonga, porém, dedicava-se principalmente ao trabalho básico nas plantações de café ou cacau, ou no processamento das colheitas, e ao mesmo tempo fazia trabalhos domésticos e cuidava dos filhos. Em resumo, a mulher teve menos convívio fora do contexto dos contingentes de serviçais, e isso reflete-se no seu uso do morfema PL.

11.7 A relevância dos processos de aquisição de segunda língua e a transferência estrutural

Por que a posição adjacente ao núcleo desempenha o papel de ponto de entrada da morfologia de PL? Por que é que esse perfil se mantém saliente nas

gramáticas das gerações subsequentes? Achamos especialmente instrutivo o caso dos tongas. Refinando as hipóteses de Baxter (2004), consideramos que os possíveis motivos para o papel fulcral da posição pré-nuclear adjacente estão ligados à natureza dos DLP disponíveis ao aprendiz africano e afrodescendente. A gama de modelos de *português* presentes no âmbito dos tongas de Monte Café incluiria, principalmente:

(i) estruturas básicas do português L1 de administradores (p.ex., DET+N) com ou sem dados de concordância variável (NARO; SCHERRE, 2007);

(ii) português L2 de adultos, inspirados em estruturas básicas do PE L1;

(iii) português L2 de adultos, com transferências estruturais das L1 africanas;

(iv) português L1 de afrodescendentes da mesma comunidade;

(v) L1 língua crioula;[24]

(vi) português L1 de afrodescendentes de fora da comunidade.[25]

Dadas as circunstâncias demográficas da roça Monte Café, parece lógico que os modelos prevalentes, em termos de número de falantes, seriam variedades de L2 e L1 de afrodescendentes. Um perfil semelhante de modelos de *português*, com predomínio de L2 e L1 de afrodescendentes, parece relevante para as décadas iniciais da Colônia Leopoldina (BAXTER; LUCCHESI, 1999; LUCCHESI; BAXTER, 2006). Consideramos de interesse especial tanto os fatores (i) e (ii), por um lado, como (iii) e (iv), por outro, para os contextos sociolinguísticos históricos dos dois dialetos em questão.

Os fatores (i) e (ii), que invocam questões de estruturação regular e básica (KLEIN; PERDUE, 1997; MATHER, 2000, 2006; SIEGEL, 2004) poderiam desempenhar um papel significativo em relação a dados da classe de morfemas funcionais (*closed class of morphemes*), D (determinante), referida em investigações recentes sobre a aquisição de PB L1 (CORRÊA; AUGUSTO; FERRARI-NETO, 2005). Esses estudos psicolinguísticos demonstram que, na aquisição da L1, o reconhecimento da variação sistemática SG/PL dos elementos funcionais D (classe de morfemas fechada) facilita a identificação do morfema de número, que é tratado como uma categoria funcional. O PL entra no sistema a partir do elemento D.

[24] Crioulo de São Tomé ou crioulo de Cabo Verde, no caso dos tongas. Dados inspirados no CST poderiam ter entrado no português no período da fundação da roça, no início da década de 1860, mas a documentação histórica aponta para mão de obra exclusivamente africana, e predominantemente angolana, já na segunda metade do século XIX. Finalmente, a presença cabo-verdiana não parece muito relevante, porque foi mínima até a década dos anos 1940.

[25] Na época colonial, o papel do tonga como intermediário e ajudante teria facilitado o contato com o português fora de Monte Café. Essas circunstâncias teriam paralelos com a Colônia Leopoldina. Nas fases de criação dos dois estabelecimentos agrícolas, haveria escravos e contratados que chegaram já falando o português L2 ou L1.

Por outro lado, considerando a presença histórica de grupos homogêneos de falantes de línguas africanas, parecem ser de interesse especial os fatores (iii) e (iv). Em sociedades multilíngues em que se tem desenvolvido uma segunda língua veicular e/ou oficial pós-colonial como L2 — por exemplo, o inglês da Singapura (BAO, 2005; SIEGEL, 2006) ou o português de Angola (INVERNO, 2005; MINGAS, 2000) e Moçambique (GONÇALVES, 1996, 2004) —,[26] há amplas evidências da influência das línguas maternas originais dessas comunidades sobre as emergentes L2/L1. Uma série de investigações recentes apontam para a importância de processos de transferência, relexificação e convergência em SLA (WINFORD, 2000), em situações de bilinguismo (SÁNCHEZ, 2006), e também na formação das línguas crioulas (LEFEBVRE, 1998; LUMSDEN, 1999; MUFWENE, 1990; SIEGEL, 1999, 2000, 2003).

Nas investigações sobre a aquisição das L2, a hipótese de *Full Transfer/Full Access* (SCHWARTZ; SPROUSE, 1994) prevê que, no início da aquisição da L2, a gramática da L1 é transferida completamente (mas sem as matrizes fonéticas dos lexemas)[27] para constituir a fase ∅, a partir da qual a gramática passa por reestruturações orientadas pelo *input* e pela GU. Por outra parte, a *Failed Functional Features Hypothesis* (HAWKINS; CHAN, 1997) prevê dificuldades na aquisição da L2 de categorias funcionais parametrizadas da língua-alvo quando a L1 do aprendiz não possui o parâmetro em questão.[28] À luz dessas investigações, parece pertinente considerar a estrutura do SN e o tratamento do número nas principais línguas africanas presentes na Colônia Leopoldina e em Monte Café.

Para o português brasileiro popular, Guy (1981a, p.300-302) propõe a hipótese de que a marcação variável da morfologia de PL no SN, atribuída a partir da esquerda do SN, poderia ter suas origens nos mecanismos de marcação de PL existentes nas línguas africanas presentes na formação da sociedade brasileira. Guy sublinha a importância numérica das línguas níger-congo dos grupos kwa, sobretudo o iorubá, e o banto, de Angola e da bacia do Congo. Aliás, ele aponta para o efeito que o contato entre essas línguas e o português teve no desenvolvimento do SN nas línguas crioulas de base portuguesa da África: no crioulo de

[26] Gonçalves (1996, 2004) apresenta dados do Português L2 de africanos moçambicanos aprendido em contexto formal a partir de um modelo padrão, que manifestam diversas reestruturações, incluindo a concordância SN PL variável.

[27] Em parte, isso faz lembrar aspectos da teoria de *code-switching* de Myers-Scotton (1997), que afirma que, num contexto multilingue, os padrões sintáticos (*syntactic frames*) de uma língua *x* podem ser incorporados em uma língua *y*, e assumidos por material fonético da língua *y*.

[28] Em essência, a FFFH também aponta para a influência da L1.

Cabo Verde (veja-se a seção seguinte) e no crioulo de São Tomé, o marcador de PL tende a ocorrer no início do NP.[29]

As línguas em questão têm em comum a caraterística de situarem um marcador de PL no início do SN, muitas vezes como um prefixo ou clítico, como é o caso das línguas banto, ou por meio de um marcador de PL ou um quantificador. Aliás, esses marcadores são comuns no início do SN, adjuntos ao classificador nominal. É o caso, por exemplo, das línguas balanta, papel e wolof, do grupo atlântico (WILLIAMSON; BLENCH, 2000), substratos do crioulo de Cabo Verde. Também é o caso das línguas da família kwa, entre as quais o iorubá e o fon, e as línguas da família banto, incluindo o kimbundo e o umbundo, todas línguas significativas na história do Brasil, e algumas delas relevantes para o caso de Monte Café (BAXTER, 2002). No iorubá, o pronome de terceira pessoa de PL *awon* é colocado antes do nome, como morfema independente marcando PL, e o iorubá não tem concordância morfossintática no SN (WILLIAMSON; BLENCH, 2000):

(3)	*omode*		*kunrin*	*giga*
	pessoa-jovem		masculino	alto
	‘O rapaz alto’			

(4)	***Awon***	*omode*	*kunrin*	*giga*
	3PL	pessoa-jovem	masculino	alto
	‘Os rapazes altos’			

Por tanto, o falante adulto de iorubá, que não tem os traços funcionais relativamente à concordância de número, iria experimentar consideráveis dificuldades na aquisição dessa concordância em português L2, mas o uso de um marcador de PL pré-nominal já estaria na sua competência.

No caso das línguas banto, a hipótese é especialmente interessante à luz dos fatos sociolinguísticos dos tongas, cuja comunidade ainda mantêm um grau de bilinguismo português-umbundo. A língua umbundo indica o plural no núcleo nominal por meio de um classificador nominal em forma de prefixo, que identifica a classe semântica do substantivo e pode também desempenhar funções dêiticas. Além disso, o umbundo tem um sistema de concordância por prefixos que se estende aos adjetivos, aos determinantes e a alguns elementos de enlace; porém a

[29] Nos crioulos do Golfo da Guiné, por exemplo no crioulo de São Tomé, a indicação de PL é muito mais influenciada pelo substrato kwa, o plural é indicado em posição pré-nuclear pelo morfema *inen*, e não há outras indicações formais de plural no SN (ALEXANDRE; HAGEMEIJER, 2007, p.40):

Inen	*funsiônariu*	*ska*	*bi*	*golo*	*pixi*	*blatu*	*ô!*
PL	funcionário	ASP	IR	procurar	peixe	barato	ENF!

‘Os funcionários estão procurando peixe barato.’

forma dos prefixos concordantes é determinada pelo classificador âncora do substantivo. Aliás, o morfema PL no classificador é a chave para a atribuição dos concordantes plurais nos outros itens no SN, pois sua forma determina a forma dos concordantes subsequentes (BAXTER, 2003, p.18):

(5) *u-kwenje* *u-wa*
CLASSIFICADOR.SG- rapaz CONCORDANTE.SG-bom
'o rapaz bom'

(6) *a-kwenje* *va-wa*
CLASSIFICADOR.PL- rapaz CONCORDANTE.PL-bom
'os rapazes bons'

É precisamente por isso que o adulto, falante de umbundo, poderia manifestar alguma sensibilidade para a indicação do PL numa posição à esquerda e adjacente ao núcleo do SN. Paralelamente, não esperaríamos que o falante de umbundo manifestasse sensibilidade à marcação de PL em posição de sufixo pós-nominal. Um outro fator relevante é que, em determinados classificadores nominais bissilábicos, é a segunda sílaba que indica o PL: *o-mbwa* '(o/um-SG) cachorro' vs. *o**lo**-mbwa* '(o-PL) cachorros'; *ovi-ndele* '(o/um-SG) branco (i.é., pessoa branca)' vs. *o**tu**-ndele* '(o-PL) brancos'. Esse fato poderia focalizar ainda mais a atenção do aprendiz na posição adjacente pré-nuclear. Segundo a FFFH, o componente morfofonológico de itens lexicais funcionais não é parametrizado, e por isso o aprendiz de L2 pode mapear material morfofonológico novo sobre esses itens (HAWKINS; CHAN, 1997, p.189). Porém, apesar de o umbundo instanciar traços funcionais de concordância de número, o falante adulto, acostumado a um sistema de concordância prefixal em fusão com classificadores, experimentaria dificuldades para assimilar concordantes PL sufixais.[30]

À luz dos antecedentes sociolinguísticos plurilíngues das duas comunidades em estudo, é evidente que os falantes de português L2 (falantes de línguas da família níger-congo) teriam motivos adequados por desenvolver os perfís indicados na Tabela 3.

11.8 O mecanismo formal de atribuição de plural

Andrade (2003) e Lopes (2001) chamam a atenção para a relevância dos trabalhos de Myers-Scotton (1997) e de Myers-Scotton e Jake (2000a, 2000b) sobre o fenômeno da mistura de estruturas em contextos de contato linguístico.

[30] Esse fenômeno no português de Angola é comentado por Inverno (2005, p.130-135).

Com base em dados provenientes de estudos de afasia bilíngue, aquisição de L2 e alternância de códigos (*code-switching*) em falantes bilíngues, Myers-Scotton (1997) explica como os lexemas de uma determinada língua podem ser inseridos nas estruturas sintáticas de uma outra língua. Nessas misturas, a estrutura sintática matriz pode conter morfemas *early system* — morfemas que são intrínsecos ao significado do quadro (*frame*) do lexema e que são gerados junto com aquele quadro.

O Modelo 4-M (MYERS-SCOTTON; JAKE, 2000a) propõe quatro classes de morfemas: uma classe de *morfemas de conteúdo* e três classes de *morfemas de sistema*. As classes de morfemas de sistema são ativadas em diferentes fases da produção linguística. Os morfemas *early-system* são ativados no léxico mental — o nível do *lemma*, na perspetiva de Levelt (1989) —, como um tipo de satélite dos morfemas de conteúdo, e atribuem a eles determinados significados essenciais, como a definição, a referência e o número. Em contrapartida, os morfemas *late system* são atribuídos por meio de requisitos estruturais. Esta distinção coincide em certa medida com a distinção gerativista de itens *semanticamente interpretáveis* sendo itens *não-interpretáveis*.

A perspetiva de Myers-Scotton é pertinente para a presente discussão. Por um lado, o modelo estipula um mecanismo de mistura que poderia servir de plataforma para o desenvolvimento inicial do perfil estrutural da presença de PL no pré-núcleo ou no núcleo em primeira posição. Por outro lado, a noção de que alguns morfemas de PL seriam mais básicos ou fulcrais, *early system*, em contraste com outros, *late system*, que seriam atribuídos pela estrutura sintática, vem ao encontro de estudos recentes sobre a morfologia de PL no crioulo de Cabo Verde e no PB (CASTRO; PRATAS, 2006; COSTA; FIGUEIREDO SILVA, 2006).

Costa e Figueiredo Silva (2006, p.35-38) propõem que a presença da morfologia flexional num determinado item em português poderia ter duas fontes: movimento sintático tipo N para o determinante (D) ou um processo morfológico.[31] Em conformidade com a hipótese da Morfologia Distribuída (*Distributed Morphology*), propõem que um determinado morfema pode ser realizado de duas maneiras: como morfema *singleton* ou como morfema *disassociated*. Este último só aparece como reflexo da estruturação sintática, depois do *spell-out*. Segundo esses autores, o morfema PL é realizado na sintaxe, no determinante, como âncora dessa informação, e a diferença essencial entre o PE e o PB está no contraste entre morfema *singleton* e *dissociated*: o PB utiliza o primeiro tipo, e o PL é prevalente no D; mas no PE, que utiliza o segundo tipo e todos os elementos capazes de levar PL, o levam como

[31] Doravante, a discussão nesta seção é feita em torno do conceito SD (sintagma determinante), explicitado na seção 11.2 deste capítulo.

reflexo da estruturação do SD.[32] Os dados dos dois dialetos considerados aqui se enquadrariam na visão do PB assumida por esses investigadores.

Ao mesmo tempo, um recente estudo do PL no SD do crioulo de Cabo Verde (CVC) nos traz outros dados muito pertinentes. Castro e Pratas (2006) adotam para esse crioulo uma análise semelhante àquela que Costa e Figueiredo Silva (2006) propõem para o PB. Inclusive, certos aspectos do padrão do PL no SD do CVC fazem lembrar determinadas tendências do PB. No CVC, quando há um elemento determinante no DP, ou um numeral, só o determinante leva a marcação de PL, mas quando não há um determinante, o PL é realizado no primeiro elemento não-D do SN (CASTRO; PRATAS, 2006, p.12).[33] Os seguintes exemplos captam essas caraterísticas (BAPTISTA, 2002, p.36-39):

PL só no determinante ou no possessivo

(7) *kes rapariga*[34]

‘aquelas moças / as moças’

(8) *nhas mininu*[35]

‘minhas crianças’

Ausência de PL na presença de numeral

(9) *oitu fidju*

‘oito filhos’

Ausência de PL em presença de quantificador

(10) *tudu rapariga*[36]

‘todas as moças’

---

[32] Para finalidades descritivas, Costa e Figueiredo Silva (2006, p.26-33) visualizam dois registros de PB (PB1 e PB2), que podem ser utilizados pelo mesmo falante: o PB1 tem plena concordância Sujeito-V, enquanto o PB2 não a tem. Porém, afirmam (2006, p.28-29) que os dois registros têm um comportamento semelhante em relação à atribuição de PL no SD: “In general, plurality is marked just on the determiner [...] either the determiner head or all the prenominal elements bear plural morphology in Brazilian portuguese”. É evidente que essa proposta simplifica demais a realidade sociolinguística brasileira, dada a presença de variedades cultas do PB que empregam a concordância ao longo do DP, no núcleo e além do núcleo, com um alto grau de consistência.

[33] Em realidade, essa descrição coincide com a configuração estrutural do SN adotada por Andrade (2003), em que os itens em <primeira posição do SN> estão na primeira posição do SN contido no DP.

[34] Baptista (2002, p.38) observa que existe a rara concordância entre D *kes* e substantivo: *kes djentis* ‘aquelas pessoas’.

[35] Quando o elemento pré-nominal é POSS, o PL pode ocorrer só no nome: *nha mininus* ‘minhas crianças’ (BAPTISTA, 2002, p.38).

[36] Compare com PB *toda menina*. Contudo, com alguns quantificadores, o substantivo pode levar PL: *txeu mininus* ‘muitos meninos’ (BAPTISTA, 2002, p.37).

PL aparece no nome quando não há elemento prenominal

(11) *Raparigas txega sedu*

'as moças chegam cedo'

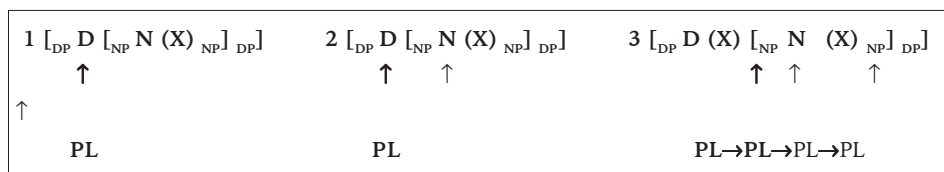
Notamos, portanto, as duas principais restrições sobre a marcação de PL: a classe de palavra, e a sua posição,[37] à semelhanca dos dois dialetos em consideração neste capítulo.

Para dar conta desses fatos, Castro e Pratas (2006) propõem que o PL presente em D é um morfema *singleton*, inserido diretamente, sendo intrínseco a D. Simultaneamente, adotam o conceito de *lowering*, uma operação morfológica pós-sintática que ocorre sob condições de adjacência e que associa morfemas abstratos uns aos outros. A categoria funcional D funciona como o eixo de controle de PL, e quando essa posição não é ocupada por um elemento D, o SN (material lexical) torna-se alvo de PL. Portanto, o PL é atribuído a partir da posição D. Isso faz lembrar o papel fulcral do PL no classificador nominal na tipologia níger-congo, e portanto no substrato níger-congo atlântico (balanta e wolof) do crioulo de Cabo Verde.

Para a descrição dos fatos dos dialetos de Helvécia e dos tongas de Monte Café, os mecanismos propostos por Castro e Pratas (2006) para a descrição do CVC e por Costa e Figueiredo Silva (2006) para a descrição do PB são bastante úteis. Seguindo essas perspectivas, a posição pré-nuclear adjacente ao núcleo é obviamente fulcral para a introdução do PL nos respetivos sistemas gramaticais. Nas faixas etárias mais avançadas, esse papel fulcral é ainda mais evidente, e a regra nesses casos é marcar o PL no pré-núcleo adjacente. Por isso, a descrição do morfema PL nessa posição como *singleton* parece lógica. Porém, é evidente que a noção de *singleton* não vai ser suficiente para dar conta dos fatos da distribuição de PL ao longo das diferentes faixas etárias. Mesmo nas faixas avançadas, as baixíssimas presenças numéricas de PL e os fortes desfavorecimentos em relação à sua inserção no núcleo em segunda posição, podem indicar o início de um processo de desenvolvimento de uma marcação como reflexo da sintaxe, inspirado em outros modelos de português, de outras gerações ou de pessoas de fora do grupo. Além disso, nas faixas etárias subsequentes, a entrada em jogo da segunda posição à esquerda do núcleo e, no caso dos tongas, a entrada em jogo do núcleo sem elemento D, e o leve acréscimo de PL nas posições mais além da segunda posição apontam para outros mecanismos de inserção de PL. Nesse sentido, postulamos que, no seu desenvolvimento, o sistema começa com um PL *singleton*, mas depois desenvolve caraterísticas de concordância parcial nos elementos mais próximos do *fulcro*: elementos do DP antes do NP (refletindo um crescimento estrutural nesse componente) e para a direita (no NP propriamente dito). Contudo, há dois fatos

[37] Baptista (2002, p.40-42) observa que a presença do morfema PL é também condicionada pelos traços [+animado], [+humano] e [-animado, +definido] no nome.

prevalentes em todas as faixas etárias, ancorando o PL: (i) o Det adjacente a N; e (ii) o primeiro item flexionável no DP. O Quadro 1 resume esse desenvolvimento.

$1\ [_{DP}\ D\ [_{NP}\ N\ (X)\ _{NP}]\ _{DP}]$	$2\ [_{DP}\ D\ [_{NP}\ N\ (X)\ _{NP}]\ _{DP}]$	$3\ [_{DP}\ D\ (X)\ [_{NP}\ N\ \ (X)\ _{NP}]\ _{DP}]$
↑	↑ ↑	↑ ↑ ↑
↑		
PL	**PL**	**PL→PL→**PL→PL

Quadro 1 - Dialetos de Helvécia e dos tongas: desenvolvimento da atribuição de PL

11.9 Conclusão

Este capítulo avaliou a relevância das propostas tradicionais relativamente à influência da configuração estrutural do SD (= SN tradicional) sobre a variável PL no dialeto afro-brasileiro de Helvécia, em comparação com o dialeto dos tongas em São Tomé. Foi demonstrado que, diferentemente das propostas recentes sobre o comportamento do PL na configuração estrutural do PB urbano, há uma forte correspondência, no dialeto de Helvécia, entre posição linear e classe gramatical. Aliás, nas análises de ambos os dialetos, o estudo revelou o papel-chave da posição pré-nuclear adjacente como âncora para a introdução da marca de PL. Nos dois dialetos, as variáveis sociais apontam para um perfil de mudança aquisicional, operando na direção da morfologia padrão. O perfil inicial da variação foi atribuído em boa parte ao substrato africano, embora não fossem descartadas outras fontes desse modelo. Finalmente, o estudo propôs um mecanismo formal misto para captar os fatos da distribuição do PL no SD dos dois dialetos: PL é introduzido inicialmente na categoria D, âncora de PL, e depois, com uma maior estruturação do SD, ele é colocado como morfema dissociado nas outras posições da estrutura. O forte perfil da âncora pré-nuclear é mantido ao longo das gerações.

O papel da estrutura pré-nuclear é uma caraterística histórica dos dois dialetos estudados e tem as suas origens nos processos de aprendizagem e transmissão linguística operantes nos contextos sociolinguísticos da Colônia Leopoldina e da grande roça Monte Café. Se houve um modelo inicial com PL variável baseado no português europeu, concordamos com as ideias de Naro e Scherre (2007) de que os africanos e os seus descendentes exageraram bastante esse modelo original. Ao mesmo tempo, à luz do que se sabe hoje em dia a respeito do papel dos processos de transferência e de reestruturação que operam durante a aquisição de L2, e de como caraterísticas de L2 podem permear o desenvolvimento de uma L1 quando essa recebe DLP de L2, parece bastante lógico assumir que o africano (ou, em outros contextos, o índio) teria influenciado o desenvolvimento do PL nos dois dialetos aqui considerados, havendo ou não variação no modelo original.

12 A concordância de gênero

*Dante Lucchesi**

A categoria gramatical do gênero apresenta-se de forma bastante diferenciada nas línguas naturais, quer em seus aspectos semânticos, quer em sua configuração morfológica.[1] Em algumas línguas, constitui apenas um marcador gramatical sem qualquer implicação no plano do significado. Em outras, relaciona-se semanticamente ao chamado gênero natural (CORBETT, 1991). Mas, mesmo nesses casos, a correspondência entre as classes mórficas e os gêneros naturais não é, na maioria das vezes, exaustiva e coerente. Na formação das línguas românicas, verifica-se a passagem de um sistema tripartido encontrado no chamado latim clássico para um sistema fundado na oposição entre o masculino e o feminino, com o desaparecimento do neutro. Esse cenário define as linhas gerais do gênero em português.

Focalizando especificamente o português afro-brasileiro no contexto da evolução histórica da língua portuguesa em cotejo com algumas línguas crioulas de base lexical portuguesa da África, este capítulo abordará os seguintes tópicos:

(i) o fortalecimento da morfologia do feminino ao longo do desenvolvimento interno da língua portuguesa, sobretudo com o surgimento de formas do feminino a partir de palavras comuns de dois gêneros;

(ii) a perda da morfologia de gênero nos processos de crioulização da língua, nos crioulos de Cabo Verde, Guiné-Bissau e São Tomé e Príncipe;

(iii) a força da marcação morfológica do gênero nos núcleos nominais em que a flexão de gênero expressa a distinção de sexo dos seres, evidenciada nos processos de descrioulização e nos processos de mudança motivados pelo contato entre línguas observados no português afro-brasileiro.

(iv) o encaixamento social que reflete os condicionamentos sócio-históricos do uso variável da regra de concordância de gênero no interior do Sintagma Nominal (SN) em uma comunidade rural afro-brasileira isolada do Estado da Bahia.

---

* Esse capítulo é uma nova versão do capítulo "A categoria gramatical do gênero: universais, mudança e crioulização" do livro *Razões e emoção*: miscelânea de estudos em homenagem a Maria Helena Mira Mateus (Lucchesi, 2003a), com o acréscimo da seção 12.5, extraída de Lucchesi (2000a).

[1] Agradeço os comentários e sugestões de Alan Baxter que contribuíram muito na finalização deste capítulo. Eventuais equívocos e omissões remanescentes são de minha inteira responsabilidade.

O isolamento de um "núcleo duro" da morfologia de gênero ligado à distinção semântica entre os sexos dos seres animados evidencia, por um lado, a motivação mais profunda da categoria do gênero, por vezes obnubilada pelo desenvolvimento histórico das línguas particulares; e, por outro, evidencia o caráter fundamental da relação entre forma e significado, enfatizada nos últimos desenvolvimentos do Programa Gerativista (CHOMSKY, 1995).

12.1 O gênero em português

As categorias gramaticais de número e gênero associam-se, no plano semântico-lexical, essencialmente ao nome e se estendem, no plano sintático, através do mecanismo da concordância, aos termos que lhe determinam o sentido: o artigo, o adjetivo, o pronome e o numeral. Em um sintagma como *as mais belas garotas*, os traços semânticos [-macho, -singular] se referem ao nome *garotas* e não ao adjetivo *belas*, ou ao artigo, ou ao advérbio, mas a sua expressão mórfica se estende a todos os constituintes flexionáveis do SN. Através do mecanismo da concordância, os artigos, adjetivos, pronomes e numerais se flexionam em conformidade com o gênero e o número do nome a que se referem. No caso do gênero, a sua indicação mórfica, na maioria das vezes, é feita exclusivamente através da concordância, na medida em que a maioria dos nomes não porta nenhuma marca formal de gênero. Em um sintagma como *as fotos coloridas*, por exemplo, o gênero feminino do substantivo *foto* é indicado apenas pela morfologia de feminino do artigo e do adjetivo. Portanto, a análise da categoria gramatical do gênero compreende, por um lado, a sua estruturação no léxico da língua e, por outro lado, a sua manifestação na estrutura sintagmática da sentença.

Na estrutura lexical dos nomes em português, a categoria gramatical do gênero, que compreende apenas dois valores, o masculino e o feminino, pode ser situada em dois planos distintos, cuja confusão conduz normalmente a uma série de incompreensões e superposições indesejáveis. A par de ser um mero índice gramatical sem qualquer implicação semântica, o gênero se apresenta também como uma categoria flexional para toda uma classe de nomes e está intimamente relacionado, nesse caso, com a distinção entre os sexos dos seres, no plano referencial.[2]

[2] Cf. Pereira (1987, p.3): "No caso do gênero, a distinção da forma e da interpretação é mais claramente reconhecível, e é frequentemente mencionada, por causa dos nítidos papéis que a diferenciação masculino/feminino desempenha, quando aplicada a substantivos animados e inanimados. Nos primeiros, a oposição se correlaciona com a distinção semântica de sexos, mas nos últimos o gênero é arbitrário, desvinculado de uma categoria semântica lexical, e apenas classifica vocábulos em classes mórficas distintas".

Na ausência de um gênero neutro para a classe dos nomes em português, as designações dos seres inanimados podem ser, tanto masculinas (como *planeta, leite, torpedo, caminhão*, etc), quanto femininas (como *porta, lente, foto, porção*, etc). Portanto, no subconjunto do léxico que contém os nomes que designam os seres inanimados, o gênero gramatical não tem qualquer relação com o conteúdo semântico do item lexical. Essa inconsistência na base semântica pode explicar as flutuações que se observam na fixação do gênero gramatical de certos nomes de seres inanimados, tanto sincrônica, quanto diacronicamente. No português contemporâneo, os gramáticos normativos (p.ex., ROCHA LIMA, 1960, p.75-76) atentam para a oscilação de gênero em palavras como *champanha, clã, diabete, gengibre, lança-perfume, cal, omoplata*, etc.[3] No plano diacrônico, são vários os casos de nomes de seres inanimados, como *planeta* e *linguagem*, que mudaram de gênero ao longo da história da língua.

Porém, no subconjunto do léxico que contém as designações dos seres animados, o gênero gramatical acompanha geralmente o chamado gênero natural, verificando-se aí uma correlação geral e consistente com o significado. Assim, são masculinas as designações dos seres do sexo masculino (*homem, pai, boi, cavalo, rei, padre*, etc), e femininas as designações dos seres do sexo feminino (*mulher, mãe, vaca, égua, rainha, madre*, etc). Contudo, vale registrar a situação especial dos chamados nomes epicenos e sobrecomuns. Os nomes epicenos possuem um único gênero e se referem a determinados animais, independentemente do sexo. A distinção entre os sexos, quando necessária, é feita com o acréscimo das palavras *macho* e *fêmea*: *Mataram uma cobra macho e capturaram um jacaré fêmea*. Já os sobrecomuns são os nomes com o traço semântico [+humano] que se comportam como os nomes epicenos. Assim, *indivíduo* e *cônjuge* são sempre masculinos, e *criança* e *testemunha* são sempre femininos, quer se refiram a indivíduos do sexo masculino, quer se refiram a indivíduos do sexo feminino. No caso dos nomes epicenos e sobrecomuns, o gênero gramatical, a exemplo do que ocorre com os nomes com o traço semântico [-animado], constitui apenas, e tão somente, um classificador mórfico, sem nenhuma implicação no plano do significado. Não há nenhuma motivação semântica, por exemplo, para *pessoa* ser feminino e *indivíduo*, masculino.

No plano mórfico, na classe dos nomes de gênero único, é igualmente difícil encontrar algum padrão completamente regular na indicação formal do gênero. Encontram-se nomes masculinos e femininos em todas as classes nominais temáticas. Assim, na classe dos nomes de tema em *-a*, temos: *cometa, poeta, poema*, que são

---

[3] Vale registrar, na recente incorporação da terminologia da informática, a solução divergente que se observa nas variedades europeia e brasileira da língua portuguesa no que tange ao termo *disquete*: feminino, em Portugal; masculino, no Brasil.

masculinos, ao lado de *mata, nora, beleza,* que são femininos; nos nomes de tema em *-e,* temos: *porrete, alfaiate, lembrete,* que são masculinos, ao lado de *estante, lebre, gente,* que são femininos; nos nomes de tema em *-o,* são exemplos de nomes masculinos: *livro, cão, trabalho,* e são exemplos de nomes femininos: *foto, mão, tribo.*

A ausência de um padrão exaustivamente coerente, tanto no plano semântico, quanto no mórfico,[4] faz com que o gênero do nome seja indicado, em português, regularmente, pelos processos sintáticos da concordância e da correlação anafórica:

> (i) masculino é todo nome a que se pode antepor o artigo *o,* ou ajuntar qualificativos terminados em *-o,* e é substituível pela palavra *ele*: *O dia claro*; *O intenso calor*; *O pano é liso, ele me agrada*.
>
> (ii) feminino é todo nome a que se pode antepor o artigo *a,* ou ajuntar qualificativos terminados em *-a,* e é substituível pela palavra *ela*: *A noite escura*; *A medonha tempestade*; *A parede é grossa, ela não cairá.* (SAID ALI, 1964, p.33).

Portanto, é através do mecanismo sintático da concordância dos artigos, bem como dos adjetivos, dos demonstrativos, dos pronomes possessivos e indefinidos, e dos numerais com o nome, núcleo do SN, que a categoria gramatical do gênero se manifesta no plano sintagmático. Desse modo, pode-se pensar o gênero como um traço inerente ao nome como um item do léxico,[5] que se expande na estrutura da sentença através do mecanismo da concordância, que compreende, além da relação de concordância entre os constituintes do SN, a relação entre um predicativo com o nome a que se refere. Entretanto, no caso dos nomes comuns de dois gêneros, os constituintes determinantes é que definem o gênero do indivíduo referido: *o meu colega / a minha colega*; *o estudante aplicado / a estudante aplicada; todos os dois jovens / todas as duas jovens*.

Por outro lado, apesar da ausência de uma morfologia regular na indicação do gênero nos nomes, observa-se uma preponderância, que remonta às origens latinas da língua, dos nomes masculinos na classe de tema em *-o* e de nomes femininos na classe de tema em *-a*. Esse padrão pode, inclusive, ser observado também entre os nomes epicenos e sobrecomuns:

(i) são sempre femininos os de tema em *-a*: *águia, baleia, borboleta* etc. (epicenos), e *criatura, pessoa, vítima* (sobrecomuns);

[4] Cf. Cunha (1981, p.131): "O gênero de um substantivo não se conhece, de regra, nem pela sua significação, nem pela sua terminação".

[5] Cf. o desenvolvimento da teoria gerativa sobre o tema: "Em *Aspects,* os traços de gênero e de classe de declinação são dados como inerentes ao substantivo, e mencionados na matriz de traços que lhe é associada no léxico" (PEREIRA, 1987, p.63).

(ii) são sempre masculinos os de tema em *-o*: *besouro*, *polvo*, *gavião*, etc. (epicenos), e *apóstolo, carrasco, verdugo*, etc. (sobrecomuns); bem como os de tema em *-e*, *-*e* ("e" teórico[6]) e os atemáticos: *antílope*, *condor* e *tatu* (epicenos), e *cônjuge, algoz* (sobrecomuns).

Pode-se associar a tal preponderância a fixação, na passagem do latim ao português, de uma flexão de gênero dos nomes que opõe uma forma flexionada de feminino, através da adjunção do morfema de feminino *-a*, à forma masculina de tema em *-o*, *-e*, *-*e* e atemática, com a supressão da vogal temática:

(i) menin*o* : menin*a*; gat*o* : gat*a*; alun*o* : alun*a*;

(ii) mestr*e* : mestr*a*; elefant*e* : elefant*a*; parent*e* : parent*a*;

(iii) leitor : leitor*a*; camponês : campones*a*;

(iv) peru : peru*a*.

Assim, podemos afirmar, em linhas gerais, que a categoria gramatical do gênero em português se estrutura de seguinte forma:

(i) entre os nomes de seres inanimados, o gênero apresenta-se apenas como classificador mórfico sem qualquer correlação sistemática no plano do significado;

(ii) no subconjunto dos nomes dos seres animados, o gênero exibe uma correlação relativamente regular com o sexo dos indivíduos, em termos semânticos.

Mas, mesmo aí, a indicação do gênero é feita de forma vária, ora através da flexão (e.g., *menino* : *menina*, *sogro* : *sogra*, etc.), ora através da oposição lexical (e.g., *genro/nora*, *frei/sóror*, *bode/cabra*) ou do processo de derivação (e.g., *barão/baronesa*, *imperador/imperatriz*, *galo/galinha*),[7] ou ainda simplesmente pela concordância dos termos determinantes no caso dos nomes de dois gêneros sem propriedade de flexão (e.g., *um estudante aplicado/uma estudante aplicada*). Para além disso, registre-se também a situação anormal dos nomes epicenos e sobrecomuns, que só possuem um gênero e designam seres animados, independentemente do seu sexo (e.g., *a cobra*, *o coiote*; *a testemunha*, *o indivíduo*). Essa situação complexa e movediça que se observa no estágio atual da língua portuguesa é muito semelhante à que se observa ao longo de sua história.

---

[6] Cf. Câmara Jr. (1970, p.76): "Os nomes terminados no singular em consoante pós-vocálica têm uma forma teórica em -e, /i/ átono final, que se deduz dos plurais. Compare-se: *feliz – felizes*, *mar – mares*, e assim por diante".

[7] Aqui consideramos definitiva a análise de Câmara Jr. (1970, p.79-80) que não reconhece em tais processos a flexão de gênero; desautorizando, assim, a análise geralmente proposta pelas gramáticas tradicionais.

12.2 O gênero na história da língua portuguesa

Na evolução histórica da língua portuguesa (sobretudo na passagem do português arcaico para o português moderno — séculos XV e XVI), verifica-se um expressivo fortalecimento da oposição masculino e feminino, com incremento da carga funcional do morfema de feminino *-a*. Assim, alguns adjetivos oriundos da segunda classe latina, que entraram na língua como invariáveis, passaram a admitir a flexão com o morfema de feminino *-a*: *portucalense* > *português* (*portuguesa*); *infante* > *infante* (*infanta*).

No português arcaico do século XIII, nomes como *senhor*,[8] *pecador*, *pastor* e *burguês* possuíam uma única forma inflexionável quanto ao gênero, que era usada tanto para o masculino, quanto para o feminino (MATTOS E SILVA, 1989, p.116); no século XIV, registram-se na mesma situação: *enganador*, *merecedor*, *vencedor*, etc. (MALER, 1964, p.27). Com a evolução da língua, todas essas formas passaram a admitir regularmente a flexão com o morfema de feminino *-a*: *senhora*, *pecadora*, *pastora*, *burguesa*; *enganadora*, *merecedora*, *vencedora*. A implementação do emprego do morfema de feminino *-a*, já atestada no português arcaico, generaliza-se a partir do século XVI e atinge especialmente os nomes e adjetivos terminados em *-or*, *-ol*, *-ês* e *-nte*: *espanhola*, *infanta* e *parenta*, a partir de *espanhol*, *infante* e *parente*; observando-se, contudo, alguma variação no seu uso até, pelo menos, o século XVIII.[9]

Pode-se dizer então que, na evolução da categoria gramatical do gênero na formação da língua portuguesa, a oposição masculino/feminino se expande e fixa uma relação mórfica através da emergência de um morfema de feminino *-a*, cuja produtividade funcional só se vem incrementando na passagem do latim para o português e na própria evolução histórica dessa língua. Portanto, qualquer mudança no sentido do enfraquecimento da oposição masculino/feminino na categoria gramatical do gênero (e de sua marcação morfológica) que se registre no panorama sócio-histórico da língua portuguesa no Brasil não pode, em princípio, ser atribuída a uma tendência observada na deriva da língua, ou a tendências já prefiguradas no seu sistema linguístico,[10] pois o que se observa aí é exatamente o contrário.

---

[8] É recorrente no *Cancioneiro da Ajuda* expressões como "Non me queredes, *mia senhor*, / fazer bem, enquant'eu viver;" (VASCONCELOS, 1904, p.230), pois *senhor* era "o título nobiliárquico dado pelos trovadores às suas damas" (VASCONCELOS, 1904, p.990).

[9] No século XVI, André de Rezende usa "a boa gente *espanhol*" (apud NUNES, 1960, p.223); e no século XVIII, atesta-se ainda "a nossa *português* casta linguagem" (apud COUTINHO, 1968, p.236), provavelmente, um conservadorismo purista.

[10] Como o fazem, por exemplo, Naro e Scherre (1993, 2007) para explicar a variação na concordância nominal e verbal no português do Brasil.

12.3 O gênero nos crioulos de base portuguesa

Nos crioulos de base lexical portuguesa falados no Arquipélago de Cabo Verde, na Guiné-Bissau, e nas Ilhas de São Tomé e Príncipe, as mudanças que afetaram a morfologia de gênero herdada do português exibem um padrão semelhante, não obstante a diferenciação no grau de reestruturação gramatical em relação à língua de superstrato que se observa entre esses crioulos.

O crioulo falado no Arquipélago de Cabo Verde, em suas duas variedades dialetais: o sotavento (falado nas ilhas situadas a sudeste) e o barlavento (falado nas ilhas situadas ao norte), é aquele que mais se aproxima estruturalmente do português.[11] Em decorrência de razões históricas e políticas, o crioulo português da Guiné-Bissau (doravante, simplesmente Crioulo) apresenta muitas semelhanças com o crioulo cabo-verdiano; mas especificamente com a variedade deste falada na capital Brava, na Ilha de Santiago (grupo sotavento).[12] Pode-se perceber, entretanto, que o Crioulo é fruto de um processo de crioulização um pouco mais intenso do que o que originou o cabo-verdiano. Já o santomense, ou *forro* (como é denominado por seus utentes[13]), falado na Ilha de São Tomé da República de São Tomé e Príncipe, é o que apresenta em sua formação o processo mais intenso de reestruturação gramatical. Na crioulização do português em São Tomé ocorreu uma redução drástica da estrutura gramatical, desencadeando uma ampla reestruturação da gramática, em que diversas estruturas se gramaticalizaram para suprir as lacunas deixadas pela erosão gramatical ocorrida na língua-alvo.

No tocante à morfologia de gênero, não se registra no crioulo cabo-verdiano (doravante CCV) praticamente nenhuma reestruturação original da gramática. Não se pode nem dizer que a morfologia de gênero do português tenha sido totalmente eliminada, como costuma ocorrer nos processos radicais de crioulização, pois observam-se atualmente no CCV algumas regras variáveis de concordância de gênero entre o nome e alguns adjetivos biformes; e, ainda que excepcionalmente, o mecanismo sintático da concordância de gênero pode atingir também o artigo indefinido. O artigo definido, o marcador por excelência do gênero em português,

---

[11] Em "um exame da forma," [...] "observa-se imediatamente que o crioulo de Cabo Verde é mais próximo ao português que os da Guiné e São Tomé" (MORAIS-BARBOSA, 1975, p.136-137) (traduzido do inglês).

[12] Cf. Wilson (1962, p.viii).

[13] A denominação vem do adjetivo *forro* 'alforriado', qualificativo dos escravos africanos libertos, segmento social que teve uma participação política destacada na história de São Tomé, vindo a constituir a elite local. O crioulo falado por esse segmento, o *forro*, distingue-se, assim, do *angolar* — crioulo falado ao sul da Ilha de São Tomé — e do *principense* — crioulo falado na Ilha do Príncipe.

não consta, a rigor, do elenco de palavras gramaticais do CCV (ALMADA, 1961, p.89; BAPTISTA, 2002, p.24-35; MEINTEL, 1975, p.210). A sua função é esporadicamente desempenhada pelo demonstrativo *kel*. Mas, mesmo nesses casos, não há a indicação mórfica do gênero, pois esta partícula — ao contrário da forma portuguesa *aquele*, da qual é derivada — não se flexiona em gênero, como se pode ver nos exemplos abaixo:

(1) *kel libru* grandi sta riba de menza.[14]

'o livro grande está em cima da mesa'

(2) *Kel kosa* ke bo dam[15]

'a coisa que você me deu'

Os nomes também não possuem flexão de gênero. Ela só se verifica em casos raros, como em *fiju/fija* 'filho/filha' (BAPTISTA, 2002, p.42-43; MEINTEL, 1975, p.221). Mas, mesmo nesses casos, a flexão encontra-se em processo de variação com a estrutura que é normalmente utilizada para marcar a distinção semântica entre os sexos nas palavras que designam seres animados: a composição lexical com os adjetivos *machu* 'macho' e *fémia* 'fêmea'. Assim, ao lado de *irmõ/irmã*, encontra-se também *irmõ machu/irmõ fêmia* 'irmão/irmã'; a exemplo do que se faz por exemplo com *kachor* 'cachorro' — sempre invariável em CCV —, *cachor machu/cachor fémia*.

O gênero pode também ser morficamente marcado no adjetivo. Mas, também aí, o processo é variável (ALMADA, 1961, p.93; BAPTISTA, 2002, p.65-70; MEINTEL, 1975, p.213), como se pode ver nos exemplos abaixo:

(3) luz sta *claru*

'a luz está clara'

(4) bu may e *noba*

'sua mãe é nova'

(5) kel mujer e *bonitu*

'aquela mulher é bonita'

(6) un *kaza feia*

'uma casa feia'

É provável que, tanto para os nomes, como para os adjetivos, a introdução dessas regras variáveis de flexão de gênero seja decorrente do processo de descrioulização ocorrido a partir de meados do século XX.

[14] Extraído de Meintel (1975, p.210).

[15] Extraído de Almada (1961, p.89).

As demais partículas que, em português, servem para marcar o gênero no SN são invariáveis em CCV. O artigo indefinido se mantém apenas na forma do masculino *un*, exibindo a propriedade de flexão de número: *uns*.[16] O demonstrativo *es* 'esse/este', a exemplo de *kel* 'aquele', só se flexiona em número, não em gênero. Os pronomes possessivos também são invariáveis — *nya* (1ª pess. sing.), *bo* ~ *bu* (2ª pess. sing.), *se* ~ *si* (3ª pess. sing.), *nos* (1ª pess. pl.), *boses* ~ *nyos* (2ª pess. pl.), *ses* (1ª pess. pl.) —, como também o são os pronomes indefinidos — *algun* 'algum/alguma', *cheu* 'muito/muita', *tudu* 'todo/toda/tudo', etc. — e os numerais — *dos* 'dois/duas', *primer* 'primeiro', etc.

Portanto, o CCV em suas variedades basiletais eliminou praticamente a morfologia de gênero oriunda do português, tendo desenvolvido um processo de gramaticalização da composição lexical com os adjetivos 'macho/fêmea' para preencher certas funções comunicativas da língua. Entretanto, em variação com este processo de reestruturação da gramática, algumas regras variáveis de flexão e concordância de gênero se manifestam nas camadas acroletais do CCV, provavelmente graças à descrioulização.

Já o Crioulo, da Guiné-Bissau, eliminou completamente a morfologia de gênero encontrada no português.[17] Em Crioulo, os nomes não se flexionam quanto ao gênero, sendo a distinção de sexo entre os seres animados marcada pelos adjetivos 'macho/fêmea': *yermon matcho/yermon femya* 'irmão/irmã' e *gatu matcho/gatu femya* 'gato/gata'. Entretanto, a exemplo do que ocorre em CCV, a gramaticalização da composição lexical com 'macho/fêmea' enfrenta, nas variedades acroletais do Crioulo, a concorrência das formas flexionadas nos nomes de parentesco: *irmon/irman* 'irmão/irmã' e *fiju/fija* 'filho/filha'. Esse processo de descrioulização não atinge os nomes com o traço [-humano], nem se verifica nas variedades basiletais (o "Crioulo fundo"), em que só se emprega a oposição lexical: *fiju matcho/fiju femya* (KIHM, 1994, p.126 e nota 2). Ainda como reflexo da descrioulização, registra-se a oposição mórfica no pronome de tratamento *nyu/nya*, empregado no registro formal e correspondente às formas portuguesas *senhor* e *senhora* (WILSON, 1962, p.17).

A indicação mórfica do gênero através do artigo (como ocorre em português) não é possível em Crioulo, pois em sua estrutura gramatical não se encontra o artigo definido, e o artigo indefinido *un* não exibe propriedade de flexão de gênero.[18]

---

[16] "A forma feminina do singular só excepcionalmente se usa, em regra para dar relevo a uma afirmação: *kel amujer tem uma kabesa!* 'aquela mulher tem uma cabeça!'" (ALMADA, 1961, p.90). Esse uso se deve provavelmente também à influência da língua de superstrato no *continuum post*-crioulo.

[17] Cf. a referência de Kihm (1994, p.126) à "notável ausência de gênero" no Crioulo.

[18] Cf. Kihm (1994, p.135): "o único item lexical que pode ser classificado [como artigo] é o artigo indefinido *un* (port. *um/uma*)".

O recurso aos adjetivos como indicadores do gênero igualmente não se aplica, pois esses também não se flexionam em gênero, sendo derivados da forma portuguesa do masculino, como se pode ver no exemplo abaixo (WILSON, 1962, p.15):

(7) i tene *kasa bonitu*

‘tem uma casa bonita’

Os demais constituintes que poderiam marcar o gênero no SN também não se flexionam em gênero. É assim com os demonstrativos *es* ‘este/esse’ e *kil* ‘aquele’, bem como com os possessivos — *nya* (1ª pess. sing.), *bo* (2ª pess. sing.), *si* (3ª pess. sing.), *no* (1ª pess. pl.), *bo* ~ *bos* (2ª pess. pl.), *se* (3ª pess. pl.) —, os numerais — *duus* ‘dois/duas’, *purmeru* ‘primeiro/a’, *segundu* ‘segundo/a’, etc. — e os pronomes indefinidos. Portanto, em Crioulo, a marcação variável do gênero é ainda mais marginal e menos significativa do que o é no CCV. Além disso, a gramaticalização do mecanismo da composição lexical com os adjetivos ‘macho/fêmea’ para expressar a distinção semântica entre os sexos está mais assentada no uso dos falantes do Crioulo do que entre os cabo-verdianos.

Em São Tomé e Príncipe, como na Guiné-Bissau, a morfologia de gênero foi quase que totalmente eliminada (HOLM, 2008, p.306). Os nomes e adjetivos são invariáveis, geralmente derivados da forma portuguesa do masculino.[19] Só uns poucos itens exibem a flexão de gênero, como *soglu/sogla* ‘sogro/sogra’, possivelmente em função do processo de descrioulização. A distinção semântica entre os sexos dos seres animados é feita, a exemplo do que ocorre no Crioulo e no cabo-verdiano, pela composição lexical; só que a oposição é feita com as palavras ‘homem/mulher’: *bwe ome/bwe mwala* ‘boi/vaca’, ou literalmente ‘*boi homem/boi mulher*’ (FERRAZ, 1979, p.60).

Não ocorre também a indicação do gênero através da concordância dos termos determinantes do nome. Uma única forma inflexionável desempenha as funções de demonstrativo e artigo definido:[20] *mwala se/ome se* ‘a/esta/essa/aquela mulher / o/este/esse/aquele homem’.[21] O artigo indefinido $u^{n}a$ é derivado da forma portuguesa do feminino *uma*, mas não se flexiona em gênero. As mesmas formas pronominais que desempenham as funções de acusativo e genitivo também não se flexionam quanto ao gênero: *mun* (1ª pess. sing.), *bo* (2ª pess. sing.), *e* (3ª pess.

[19] Vale registrar, como exceção, *mina* (< ptg. *menina*) ‘filho/filha’.

[20] O estatuto do *se* no santomense é um tema controverso. Essa visão de que ele atua como demonstrativo e artigo definido, adotada por Ferraz (1979), Lucchesi (1993) e Schang (2000) é contestada por Alexandre e Hagemeijer (2007). Para esses últimos, o *se* seria um marcador do traço [+específico] do SN. Apesar do desenvolvimento de um artigo definido a partir de um demonstrativo ser um fato bastante documentado em muitas línguas, eles argumentam que o *se* não pode ser tratado como artigo, porque ele pode coocorrer com outros marcadores de definitude.

[21] Note-se que, ao contrário do que ocorre na língua lexicaficadora, o determinante *se* (< ptg. *esse*) se coloca à direita do nome núcleo, uma mudança na ordem que Ferraz (1979) atribui às influências das línguas de substrato.

sing.), *nõ* (1ª pess. pl.), *nanse* (2ª pess. pl.), *inen* (3ª pess. pl.). Do mesmo modo, são inflexionáveis em gênero os numerais — e. g., *dosu* 'dois/duas' — e os pronomes indefinidos (cf. exemplos abaixo):

(8) mwala mun luma *tudo* kwa se

'minha mulher arrumou todas as coisas'

(9) e kompla *otlo* bola

'ele comprou outra bola'

(10) ome se ka te *mõtchi* mwala

'aquele homem tem muitas mulheres'

Portanto, em santomense, a morfologia de gênero foi drasticamente reduzida, verificando-se apenas alguns resquícios dessa categoria gramatical na flexão dos nomes. O mecanismo sintático da concordância através da flexão em gênero dos termos determinantes do nome foi completamente eliminado. Para expressar a distinção semântica de sexo entre os seres animados ocorreu a gramaticalização da composição lexical dos nomes 'homem/mulher' com o nome núcleo do SN.

12.4 O gênero no português afro-brasileiro

Na maioria das variedades linguísticas do português do Brasil, a variação na concordância de número no SN é bem significativa, chegando a ocorrer entre as camadas mais cultas e estando amplamente generalizada entre as variedades populares da zona rural, onde o plural no SN é quase que só marcado no determinante (SCHERRE, 1994). Entretanto, a variação na concordância de gênero não é comum entre as variedades rurais do português. Só se registra um nível significativo de variação na concordância de gênero em certas comunidades rurais que passaram por um amplo e profundo contato linguístico em sua história. Nesse sentido, é pertinente a presença dessa variação em materiais provenientes de outras comunidades afro-brasileiras como, por exemplo, Mata Cavalos, em Mato Grosso (VOGT; FRY, 1985, p.115); Calunga, em Goiás (BAIOCCHI, 1991); e Vale do Ribeira, no Estado de São Paulo (CARENO 1991, p.207-208). Assim, desde a perspectiva da transmissão irregular, é bem plausível que essa variável figurasse no português afro-brasileiro em épocas passadas. A variação na concordância de gênero em variedades de segunda língua é notória, persistindo mesmo no desempenho de falantes com competências avançadas na língua-alvo. Também se registra a variação na concordância de gênero em variedades vernáculas e de português como segunda língua na África (BAXTER, 2002, p.17-18; GONÇALVES; STROUD, 1998, p.106-109; INVERNO, 2005, p.137; MENDES, 1985, p.149; MINGAS, 2000, p.70).

Portanto, há um conjunto de evidências que fundamentam a hipótese de que a variação na concordância de gênero tenha sido geral no processo de formação do português afro-brasileiro, nas comunidades em que o grau de contato linguístico foi mais intenso.[22]

Essa relação histórica entre a variação na concordância de gênero e o processo de transmissão linguística irregular se reforça com o fato de que a perda das marcas flexionais através das quais se realiza o mecanismo sintático de concordância de gênero e número é muito geral nos processos de crioulização.[23] Como se viu na seção anterior, o mecanismo sintático da concordância de gênero foi profundamente afetado nos crioulos de base portuguesa da África — o mesmo se pode dizer acerca da concordância nominal de número. Para esses casos, é inconteste a relação histórica dessas mudanças com o processo de transmissão linguística irregular que está na origem dessas línguas crioulas. Como esse processo teria interferido na evolução histórica do português brasileiro, sobretudo em suas variedades mais afetadas pelo contato entre línguas, é o que se buscará demonstrar aqui, através do paralelo que se pode estabelecer entre o padrão observado nas mudanças que afetaram a categoria gramatical do gênero nas línguas crioulas e o padrão observado na comunidade rural afro-brasileira de Helvécia, situada no Extremo Sul da Bahia.[24]

Com base numa amostra de fala vernácula colhida junto a dezoito habitantes dessa comunidade, foi feita uma análise da variação na concordância de gênero no interior do SN, de acordo com os pressupostos teóricos da Sociolinguística Variacionista (WEINREICH; LABOV; HERZOG, 2006; LABOV, 1972a, 1982, 1994), com recursos a alguns elementos da Teoria da Gramática (CHOMSKY, 1981, 1986, 1993, 1995).[25]

A variação na concordância de gênero no interior do SN, na comunidade de fala de Helvécia, pode ser assim exemplificada:

1. Valor positivo – realização da concordância de gênero no SN

(11) Eu comprei *essa* casa aí *na* rua.

(12) Fiquei no soro durante *a* noite *toda*.

2. Valor negativo – não realização da concordância de gênero no SN

(13) Ele era *uma* pessoa muito *querido*.

(14) foi vendeno *tudo essas* madeira,

(15) se tirá o *certidão* de óbito do meu pai,

[22] É de notar-se que há evidências dessa variação também em dados do português L2 de comunidades ameríndias (MATTOS E SILVA, 1989, p.103-108; EMMERICH, 1992, p.85-86).

[23] Sobre isso, veja-se também: Bickerton (1988), Holm (1988) e Winford (2003, p.322-323).

[24] Para mais dados sobre a comunidade de Helvécia, veja-se a seção 2.2. do segundo capítulo deste livro.

[25] Essa análise é plenamente desenvolvida em Lucchesi (2000a).

(16) Muitas veze, 'duece um *pessoa*, num tem ambulança,

O nível de variação que se observa hoje em Helvécia é bastante reduzido, e os casos de ausência de concordância de gênero no SN ficam em torno de cinco por cento do total de ocorrências eliciadas. Entretanto, ainda se observa na comunidade de fala um quadro de variação estruturada, em que o fenômeno variável se encontra encaixado na estrutura linguística do dialeto, podendo-se estabelecer correlações sistemáticas nos diversos contextos linguísticos. Os principais fatores linguísticos condicionadores da variação na concordância de gênero podem ser, então, reunidos em três grandes princípios:

1. *princípio da simplicidade*: o mecanismo acessório da concordância se aplica mais nas estruturas mais simples (sobretudo nos SNs formados apenas pelo núcleo nominal precedido por um determinante: Det + Nome — cf. exemplo 17, abaixo) do que nas estruturas mais complexas (cf. exemplos 18 e 19, abaixo):

(17) eu falo *a* verdade

(18) chega lá é *um* maior confusão pá esses médico atendê a gente

(19) qualqué *um* parte que o sinhô me pergunta pra mim fazê, eu insino o sinhô.

2. *princípio da integração*: a regra de concordância é favorecida pelas estruturas mais integradas, de modo que os determinantes e modificadores à esquerda do nome tendem a receber mais marcas de feminino (cf. exemplos 20 e 21, abaixo) do que os modificadores em estruturas de adjunção à direita do nome (cf. exemplos 22 e 23, abaixo):

(20) Não, trabalho n*a minha* terra mesmo.

(21) *Verdadêra* terra de nós é esse lá.

(22) E, aliás, num tem cobra *ninhum* que num é braba.

(23) Ah, é... é coisa muito *bom*!

3. *princípio da saliência*: quanto mais fortes forem as marcas mórficas no nome núcleo, maior será a probabilidade de aplicação da regra de concordância. Desse modo, os SNs cujo núcleo é um nome com propriedade de flexão (cf. exemplo 24, abaixo) tendem a apresentar a concordância plena de gênero numa frequência mais elevada que os SNs cujo núcleo é um nome sem a propriedade de flexão (cf. exemplo 25, abaixo); esses, por sua vez, apresentam um nível de concordância superior aos SNs cujo núcleo é um substantivo comum de dois gêneros (cf. exemplo 26, abaixo).[26]

(24) e adepois, juntô com *a sogra*.

[26] Observe-se que esse fato contraria completamente uma lógica funcional da língua, já que, com os nomes comuns de dois gêneros como *caçula*, *estudante*, etc., a concordância é o mecanismo crucial para se prover a informação referencial acerca do sexo desses indivíduos: *o estudante aplicado/a estudante aplicada*.

(25) cada um tem *um* natureza.

(26) Dipôs ela teve Alicia que é *o* caçula.

Da mesma forma, entre os nomes sem propriedade de flexão de gênero, a concordância é mais frequente entre os nomes de tema em -*a* (cf. exemplo 27, abaixo) do que entre os nomes de tema em -*o* (cf. exemplo 28, abaixo). Esse fato se deve provavelmente à homonímia entre a vogal temática e o morfema -*a* de feminino, assim como à larga predominância de nomes femininos na classe dos nomes de tema em -*a* e à larga predominância dos nomes masculinos na classe dos nomes de tema em -*o*, no português (cf. seção 11.2, supra).

(27) Às vez, puxa um poquinho d*a* perna.

(28) Esse é o moço [...] que tirô *um* foto aí de Vanda.

O princípio da saliência decorre de um fato fundamental para a definição do padrão de marcação do gênero em Helvécia: a flexão de gênero é praticamente categórica nos núcleos nominais que exibem propriedade de flexão, mesmo quando todos os outros constituintes do SN não se flexionam, como nos exemplos:

(29) Agora *o* sinhor*a* lembra de tudo.

(30) *Os* minin*a* d'agora *tudo* achô coisa bom.

Esse padrão estrutural de variação na concordância de gênero no interior do SN que se observa na comunidade de fala de Helvécia decorre de um processo histórico marcado pelo contato entre línguas que pode ser escrutinado a partir das evidências do encaixamento social desse processo de variação, partindo-se da perspectiva do que se tem denominado de *tempo aparente* (LABOV, 1981; LUCCHESI, 2001b, entre outros). A análise do condicionamento sócio-histórico do fenômeno variável em foco será feita na próxima seção.

12.5 Condicionamentos sócio-históricos da variação na concordância de gênero no português afro-brasileiro

As informações disponíveis acerca da sócio-história de Helvécia apontam para um cenário altamente favorável ao processo de transmissão linguística irregular ocorrido durante o século XIX, com a implantação, na região, de empreendimentos agroexportadores de café por colonos suíços, franceses e alemães que empregavam largamente a mão de obra de escravos africanos (cf. seção 2.2.2 do segundo capítulo deste livro). Portanto, a caracterização sócio-histórica do processo de mudança em progresso na comunidade de fala de Helvécia no sentido da fixação do uso da regra de concordância de gênero em um nível praticamente categórico estrutura-

se a partir da hipótese de que, durante a formação da comunidade de fala de Helvécia, entre meados do século XIX até as primeiras décadas do século XX, houve uma redução do uso da regra de concordância de gênero, em função da aquisição precária do português por parte dos escravos africanos e da nativização desse modelo defectivo de segunda língua entre os descendentes desses escravos.[27] A partir de então, o uso da regra de concordância de gênero vai se implementando na comunidade, à medida que aumenta sua relação com o mundo exterior, com o deslocamento de indivíduos para fora e para dentro da comunidade, a influência crescente dos meios de comunicação de massa e dos programas de educação pública, etc. A análise dos resultados quantitativos das variáveis sociais fornece a fundamentação empírica para a constituição desse cenário sociolinguístico em que o nivelamento linguístico, a partir da difusão dos padrões linguísticos urbanos, elimina as marcas mais características do processo anterior de transmissão linguística irregular.

12.5.1 A variável faixa etária

De acordo com os parâmetros adotados para as análises em tempo aparente, o ponto de partida para a caracterização dos processos de mudança em curso é fornecido pelos resultados da variável faixa etária (LUCCHESI, 2001b). Os informantes que compõem a amostra de fala analisada estão equitativamente distribuídos em três faixas etárias: de 20 a 40 anos (faixa 1); de 41 a 60 anos (faixa 2); e de mais de 60 anos (faixa 3). Desse modo, fica garantida a possibilidade de uma projeção para um período correspondente a três gerações, já que nas comunidades rurais brasileiras os indivíduos tendem a constituir família e/ou procriar no final da adolescência (entre 16 e 20 anos). A frequência de uso da regra de concordância de gênero entre essas gerações da comunidade pode ser vista na Tabela 1, bem como em sua representação gráfica, feita com base nos pesos relativos.

Tabela 1 - Uso da regra de concordância de gênero, em Helvécia-BA, segundo a variável faixa etária

FAIXA ETÁRIA	Nº DE OCORRÊNCIAS/TOTAL	FREQUÊNCIA	PESO RELATIVO
Faixa 1 (20-40 anos)	1.365/1.396	98%	.65
Faixa 2 (40-60 anos)	1.286/1.331	97%	.57
Faixa 3 (mais de 60 anos)	1.193/1.296	92%	.28
Total	3.844/4.023	96%	-

Nível de significância: .027.

[27] No que se denomina aqui de *transmissão linguística irregular* (cf. capítulo 3).

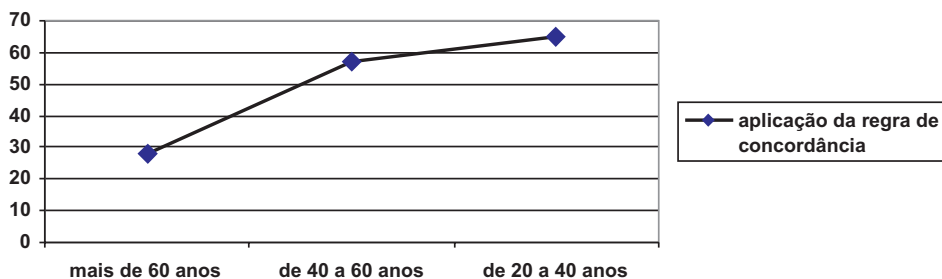



Figura 1 - Aplicação da regra de concordância de gênero no interior do SN, em Helvécia-BA, segundo a variável faixa etária (pesos relativos)

O padrão observado indica um processo de mudança em progresso, já que o ápice da variante inovadora, a aplicação da regra de concordância de gênero, encontra-se entre os falantes mais jovens, que registram percentuais de aplicação da regra da ordem de 98% (peso relativo de .65), enquanto os mais velhos apresentam um maior grau de variação: 92% de frequência de aplicação da regra, com um peso relativo de .28. Nota-se ainda, considerando sobretudo o valor dos pesos relativos, uma maior semelhança entre os padrões de uso linguístico dos falantes das faixas 1 e 2, em oposição a um padrão de variação um pouco mais acentuado entre os falantes da faixa 3. Isso indica que o processo de mudança de aquisição da regra de concordância teria se acelerado nas duas últimas gerações. De qualquer forma, o nível de aplicação da regra é elevado mesmo entre os falantes com mais de 60 anos.

Colocam-se, assim, duas hipóteses alternativas: (i) a variação na marcação do gênero não foi muito intensa no processo de transmissão linguística irregular ocorrido no século XIX; (ii) a sua recomposição foi muito rápida. De acordo com a hipótese clássica da abordagem em tempo aparente (NARO, 2003), os indivíduos com mais de 60 anos fixaram a sua gramática de uso da língua no final da adolescência, ou seja, entre 1924 e 1944, para quem tinha, em 1994, entre 65 e 85 anos. Mesmo considerando a possibilidade de esses indivíduos terem aumentado a frequência de aplicação da regra ao longo da sua vida, conforme argumentação de Lucchesi (2001b, p.137), o nível de aplicação da regra já seria relativamente alto na época. Desse modo, pode-se pensar favoravelmente à primeira hipótese, de que a marcação do gênero no SN foi menos afetada pelo processo de transmissão linguística irregular na formação das variedades populares do português no Brasil do que a marcação do número (LUCCHESI, 2000a, p.19-21).

De qualquer forma, o uso da regra de concordância de gênero foi intensificado/incrementado na passagem da década de 1950 para 1960, na comunidade de Helvécia, o que se reflete na diferença que separa a faixa 3 das demais. Observa-se nesse período (sobretudo na década de 1960 e início da década de 1970): o aumento dos movimentos migratórios (o que implica a saída e o retorno de trabalhadores da

comunidade que vão buscar serviço nos grandes centros urbanos), a abertura de rodovias e a implantação das grandes empresas agroindustriais de celulose na região, e a chegada na localidade dos meios de comunicação de massa (primeiramente o rádio, depois a televisão) e dos programas de educação pública.

Portanto, a ação desses fatores socioeconômicos e culturais criou as condições para que ocorresse o nivelamento linguístico que tende a eliminar uma marca mais característica da história crioulizante do dialeto, em função do padrão de realização da concordância de gênero, que é o mais generalizado no português do Brasil e que conta com o respaldo institucional normativo. Essa influência do padrão urbano atua livremente na comunidade, já que não se encontra um sentimento de reação como o observado por Labov na localidade de Martha's Vineyard, em 1963. Há, no geral, uma visão positiva acerca dos benefícios trazidos com o progresso, a melhoria nas condições de transporte, no atendimento médico e a possibilidade de integração cultural proporcionada pelos meios de comunicação de massa. Os mais jovens são mais atingidos por esse sentimento, buscando de uma maneira mais ativa a sua integração no novo *modus vivendi*. Os mais velhos também se mostram receptivos, afirmando que a vida de hoje é menos dura do que a de antigamente. Destacam, sobretudo, a facilidade de transporte dos doentes e a possibilidade, mesmo que precária, de atendimento no sistema de saúde pública. Todos, jovens e velhos, admitem que o parto assistido nos hospitais é mais seguro do que o feito em casa, e as parteiras bem como o recurso às plantas medicinais e rituais de cura (denominados *simpatias*) estão se tornando coisas do passado.

Os mais velhos afetam um grande reconhecimento em relação ao sistema de previdência pública. A aposentadoria rural, no valor de um salário mínimo, desempenha um importante papel na economia local, garantindo não apenas o sustento do idoso, mas contribuindo de maneira decisiva para o sustento de toda a família. A aposentadoria garante assim o respeito ao idoso, que deixa de ser um fardo e torna-se uma importante fonte da renda familiar.

Mesmo a implantação das grandes plantações de eucalipto para a produção de celulose, que alterou a distribuição de terra na região e trouxe prejuízos para os pequenos agricultores, é encarada com resignação. Essas plantações de eucalipto e a extração madeireira desarticularam bastante a cultura primitiva da região. A caça desapareceu com a erradicação das matas, e a pesca nos rios e represas é, em muitos casos, vedada pelas empresas proprietárias das terras. As famílias que venderam as suas terras para as empresas agroindustriais se desestruturaram e normalmente se encontram numa situação econômica pior do que quando cultivavam sua pequena propriedade. Os que se mantêm na vizinhança das grandes plantações sofrem com o esgotamento da terra, devido à extenuante cultura do eucalipto; porém não há consenso entre os moradores se o eucalipto, de fato, seca

a terra para garantir o seu crescimento vertiginoso. E ainda há quem reconheça que essas empresas acabam por ter um papel social importante, porque empregam aqueles que perderam a sua fonte de renda quando lhes venderam sua terra.

Por fim, os pequenos agricultores da faixa 1, que estão assumindo uma posição de liderança na comunidade, estão voltados apenas para os benefícios que o desenvolvimento tecnológico pode trazer para as suas plantações (a possibilidade de utilização de tratores, fertilizantes e agrotóxicos), tendo pouca consciência acerca dos seus efeitos colaterais — como os casos de intoxicação de agricultores (que podem ser fatais) e prejuízos ecológicos. Entre os mais conscientes, detecta-se apenas uma postura mais reivindicativa frente aos programas sociais de saúde e educação.

Nesse cenário sociocultural, os padrões linguísticos que adentram a comunidade são recebidos como os frutos do progresso, ora com adesão, ora com resignação. O abandono dos usos linguísticos crioulizantes, como a não concordância de gênero, se dá de uma forma pouco consciente. Ou seja, não há um julgamento muito explícito das variantes linguísticas no plano da avaliação social. De um lado, a não realização da concordância não chega a constituir um estigma; talvez, entre os mais jovens, possa ser abertamente rejeitada. Mas, no geral, os falantes não possuem esse grau de julgamento normativo.[28] Por outro lado, a realização da concordância de gênero pode assumir as feições de um marcador social, no bojo de uma visão positiva de integração no mundo exterior; nesses casos, pode-se supor como possível uma variação estilística não muito acentuada, na qual o falante busca realizar a concordância na proporção de sua maior atenção ao ato de fala.[29] De qualquer forma, seja num nível mais consciente, seja num nível menos consciente, o prestígio está associado à realização da concordância, e o movimento da mudança, tanto do ponto de vista social, quanto no nível de consciência dos falantes, pode ser definido como de cima para baixo e de fora para dentro da comunidade (LABOV, 1972a, 1982, 1994).

Portanto, o nivelamento linguístico nada mais é do que um reflexo da integração positiva, ou negativa, mas inexorável, da comunidade ao sistema capitalista rural brasileiro. Assim sendo, as demais variáveis sociais serão analisadas a partir dessa visão de que a mudança advém da influência externa e está intimamente ligada à integração da comunidade no mundo exterior. Se as variáveis estada fora da comunidade e nível de escolaridade expressam diretamente esse processo, a variável sexo também pode ser analisada segundo esse prisma.

---

[28] As condições em que foi realizada a pesquisa de campo, os objetivos primários da recolha e principalmente as limitações impostas pelas precárias condições de vida dos informantes não permitiu realizar testes de avaliação subjetiva, de modo que essas considerações são feitas com as impressões colhidas durante a pesquisa de campo na comunidade, em 1994.

[29] Para essa tipologia do julgamento social das variantes linguísticas, veja-se Labov (1972a) e Lucchesi (1994, 2002b).

12.5.2 A variável sexo

Diferentemente do que ocorre nos grandes centros urbanos das sociedades industrializadas, em que as mulheres lideram as mudanças em direção ao padrão (CHAMBERS; TRUDGILL, 1980), são os homens que estão à frente da fixação da concordância de gênero em Helvécia; o que se coaduna mais com o padrão observado nas comunidades rurais e na periferia das grandes cidades do Brasil (LUCCHESI, 2007). Os resultados da variável sexo são apresentados na Tabela 2.

Tabela 2 - Uso da regra de concordância de gênero, em Helvécia-BA, segundo a variável sexo

SEXO DO INFORMANTE	Nº DE OCORRÊNCIAS/TOTAL	FREQUÊNCIA	PESO RELATIVO
Homem	1.861/1.944	96%	.56
Mulher	1.983/2.079	95%	.45
Total	3.844/4.023	96%	-

Nível de significância: .027.

Como se pode observar, os homens superam as mulheres no uso da forma padrão, com peso relativo de .56 para a realização da regra entre os homens contra .45 entre as mulheres. O parâmetro do contato com o mundo exterior possibilita a compreensão do fato. Os homens, no contexto cultural da comunidade, tendem a ter mais contato com o mundo exterior. São os homens que mais vão à feira vender produtos da lavoura, e foram eles que trabalharam na ferrovia Bahia-Minas e atualmente trabalham nas grandes plantações de eucalipto. Essa maior integração do homem no processo produtivo e na vida social o coloca numa situação de maior exposição aos padrões linguísticos adventícios e numa posição que favorece mais a sua assimilação. Já as mulheres, envolvidas com a criação dos filhos e com as tarefas domésticas, ficam mais restritas à vida isolada da pequena propriedade, não obstante também trabalharem na lavoura e também levarem os produtos para a feira. Circunscritas ao universo doméstico, acabam por conservar mais os primitivos usos linguísticos, mais diretamente marcados pelo processo de transmissão linguística irregular.

Padrões semelhantes em relação à variável sexo foram encontrados por Rodrigues (1992, p.169), em estudo sobre a concordância verbal na periferia da cidade de São Paulo, e por Bortoni-Ricardo (1985) para os falantes *rurbanos* de Brazlândia, cidade satélite de Brasília. Se, no universo da norma culta, dos falantes das classes médias urbanas, as mulheres têm-se mostrado mais sensíveis às formas do padrão, o quadro entre os segmentos populares é diverso, tanto na periferia das grandes cidades, quanto nas comunidades rurais do interior do país. Nesse universo, a maior inserção dos homens no mercado de trabalho vai fazer com que esses liderem os processos de mudança em favor do padrão urbano culto.

12.5.3 A variável estada fora da comunidade

Considerando que a mudança se dá pela importação de padrões linguísticos e culturais externos, é natural que os membros da comunidade que viveram um período fora da comunidade estejam na vanguarda do processo. Nesse sentido, os informantes foram divididos entre aqueles que passaram pelo menos seis meses fora da comunidade e aqueles que nunca saíram da localidade ou que só se ausentaram por curtos períodos de tempo. Essa estada fora se dá normalmente em função da busca de trabalho nos grandes centros urbanos, ou mesmo do trabalho agrícola em regiões mais desenvolvidas. No caso da emigração para os centros urbanos, os homens se empregam normalmente como trabalhadores braçais na construção civil ou como vigias ou empregados subalternos de empresas de qualquer ramo, enquanto as mulheres, via de regra, como empregadas domésticas. Os resultados da variável estada fora da comunidade estão apresentados na Tabela 3.

Tabela 3 - Uso da regra de concordância de gênero, em Helvécia-BA, segundo a variável estada fora da comunidade

ESTADA FORADA COMUNIDADE	Nº DE OCORRÊNCIAS/TOTAL	FREQUÊNCIA	PESO RELATIVO
Sim	1.511/1.566	96%	.56
Não	2.333/2.457	95%	.46
Total	3.844/4.023	96%	-

Nível de significância: .027.

Os resultados demonstram, de forma consistente, que os membros da comunidade que viveram um período fora tendem a realizar mais a concordância de gênero do que aqueles que nunca se ausentaram por um lapso de tempo significativo, com peso relativo de .56 contra .46. Entretanto, ao que tudo indica, o efeito dessa variável parece ser mais significativo entre os falantes mais jovens. Isso se explica pela mudança no espectro da emigração ocorrida nas últimas décadas. Até a década de 1960, os trabalhadores que deixavam a localidade se empregavam geralmente na própria região, fosse no trabalho agrícola, fosse em atividades urbanas nos pequenos centros urbanos regionais. Os informantes que viveram nos grandes centros do Sul do país (Rio de Janeiro e São Paulo, principalmente) concentram-se na faixa etária de 20 a 40 anos, muitos dos quais, mulheres que haviam se empregado no trabalho doméstico, "em casa de família". Nesses casos, a assimilação da pronúncia, das gírias, dos padrões linguísticos urbanos como um todo é bastante notável. Pode-se pensar que, até o início da década de 1970, os usos linguísticos que os membros da comunidade encontravam nas regiões vizinhas ou nos pequenos centros urbanos regionais não deveriam divergir muito dos seus ou, pelo menos, não deveriam provocar um maior efeito. A influência maior se deu, então, a partir do contato direto com o padrão dos grandes centros urbanos, em função até das

atuais condições de maior difusão linguística e cultural. É nesse contexto que se pode compreender melhor o nivelamento linguístico de todas as regiões do país a partir da ação reitora das grandes metrópoles brasileiras.

12.5.4 A variável nível de escolaridade

O fato de o processo de mudança em curso se dar através da difusão de modelos normativos coloca em realce a variável nível de escolaridade. Entretanto, na zona rural em que se encontra a maioria dos membros da comunidade de Helvécia, praticamente não há alfabetizados funcionais plenos. No geral, os membros da comunidade têm acesso a apenas um ano de escolaridade. Diante disso, os informantes se dividem entre aqueles que haviam tido qualquer experiência no sistema de educação pública e aqueles que não tiveram qualquer experiência de alfabetização. Desse contato mínimo com a escola, muitas vezes, só resulta a capacidade de "assinar o nome", o que tem um efeito mais simbólico no plano da cidadania do que funcional, pois o indivíduo, sendo capaz de desenhar o próprio nome, pode receber seus documentos e ter acesso a certos serviços públicos sem se expor a uma situação vexatória, decorrente da revelação do seu pleno analfabetismo num mundo regido pelo registro escrito. Porém, mesmo diante da fragilidade dos resultados dos programas de educação pública, a expectativa é a de que os semianalfabetos apresentem uma melhor *performance* no uso das formas do padrão do que os analfabetos plenos.

Os resultados das frequências absolutas corresponderam a essa expectativa, com os semianalfabetos realizando plenamente a concordância de gênero em 98% dos casos, enquanto os analfabetos a realizavam em 94% dos casos. Contudo, esse grupo de fatores não foi selecionado pelo Programa das Regras Variáveis. Ao que tudo indica, o problema está numa incontornável superposição na distribuição das variáveis na amostra analisada. Enquanto na faixa 1 (de 20 a 40 anos), cinco dos seis informantes eram escolarizados; na faixa 2 (41 a 60 anos), só dois informantes eram escolarizados; e na faixa 3 (de informantes com mais de 60 anos), todos os informantes eram analfabetos. Além disso, os dois informantes escolarizados da faixa 2 eram homens. Com efeito, no universo dos pequenos agricultores da localidade de Helvécia, praticamente não se encontram indivíduos escolarizados com mais de 60 anos. A pobreza, o isolamento e o abandono da comunidade até a primeira metade do século XX não permitiram que esses indivíduos tivessem acesso a qualquer forma de educação formal. Só os membros das novas gerações foram beneficiados pelos programas públicos de alfabetização implementados de forma mais intensa a partir da década de 1970. E o fato de os dois únicos informantes semialfabetizados da faixa 2 serem homens e o único

informante analfabeto da faixa 1 ser uma mulher só confirma o que já foi dito aqui acerca da posição privilegiada dos homens em relação às mulheres, no que diz respeito à participação na vida pública.

Diante disso, o fator escolarização pode ser considerado de uma forma indireta. Os falantes mais jovens e do sexo masculino realizam mais a concordância porque, entre outras coisas, tiveram (mais) acesso aos programas de educação pública. Contudo, não se deve deixar de ter em mente a precariedade de tais programas, antes de fazer qualquer afirmação sobre os efeitos linguísticos da escolarização no meio rural. Em muitos casos, os professores recrutados nesses programas são muito mal preparados; e, sob a ótica do padrão linguístico normativo, praticamente não se diferenciam dos seus alunos.

12.5.5 Conclusão da análise do encaixamento social

Os resultados das variáveis sociais se mostraram bastante consistentes com a avaliação do quadro de variação observado na comunidade de fala de Helvécia como sendo o reflexo sincrônico de um desenvolvimento diacrônico da gramática da comunidade no sentido da fixação da regra de concordância de gênero num nível praticamente categórico. A distribuição pelas três faixas etárias representa um padrão ascendente de aplicação da regra de concordância de gênero. Esse padrão ascendente geracional reflete um processo de mudança em curso na comunidade que é liderado pelos homens com alguma escolaridade que viveram pelo menos seis meses fora da comunidade, trabalhando nos grandes centros urbanos. Esse cenário sociolinguístico configura um processo de mudança de cima para baixo, no qual o padrão urbano tem se difundido para todas as regiões do interior do país, através sobretudo da ação maciça dos meios de comunicação de massa, com a contribuição da ampliação do sistema de educação pública e do deslocamento populacional. E tal processo de nivelamento linguístico vai apagando os resquícios do processo de transmissão linguística irregular desencadeado nas situações de multilinguismo que caracteriza o interior do país até o século XIX. Esse é o caso da variação na concordância de gênero no interior do SN que tende em breve a desaparecer dos padrões de comportamento linguístico da comunidade.

12.6 Conclusão

A análise reuniu, assim, evidências empíricas de que, em estágios anteriores da evolução diacrônica do dialeto de Helvécia, teria ocorrido um processo mais intenso de variação na concordância de gênero provocado pelo processo de

transmissão linguística irregular, desencadeado pela aquisição imperfeita do português pela maioria de falantes africanos trazidos da África para a região do dialeto, no século XIX, e se implementado com a nativização desse modelo defectivo de português nas gerações seguintes dos descendentes *crioulos* desses escravos africanos. E o aspecto mais decisivo para definir o caráter crioulizante da variação em Helvécia é o fato de o padrão estrutural da variação nessa comunidade ser o mesmo dos processos de mudança que afetaram a concordância de gênero do português na formação das línguas crioulas na África.

Nos crioulos de base lexical portuguesa da África, a erosão da morfologia flexional do nome no que diz respeito à categoria gramatical do gênero afetou mais o mecanismo morfossintático da concordância do que a flexão mórfica do gênero nos núcleos nominais que designam os seres animados. Assim, o mecanismo morfossintático da concordância de gênero, que no português serve para indicar o gênero gramatical exaustivamente no conjunto de nomes da língua (na maioria dos casos, sem qualquer relevância no plano da informação referencial), foi praticamente eliminado; só se mantendo marginalmente nas variedades acroletais do crioulo cabo-verdiano, aquele que apresenta o menor grau de crioulização. Já a flexão de gênero nos nomes dos seres animados conservou-se, por possuir o valor referencial, mesmo que de forma variável, em todos os crioulos portugueses da África, inclusive no santomense, que exibe o maior grau de crioulização. Portanto, no que se refere à categoria gramatical do gênero, a estrutura linguística que tende a oferecer a maior resistência à erosão gramatical ocorrida durante o processo de transmissão linguística irregular é a marcação do gênero nos núcleos nominais que designam os seres animados, nos casos em que a indicação referencial do sexo dos indivíduos é relevante.

A análise da variação na concordância de gênero no SN em Helvécia revelou que o processo de variação e mudança nesse dialeto afro-brasileiro segue o mesmo padrão observado nas línguas crioulas da África, observando-se apenas uma diferença no grau de intensidade das mudanças e variação em cada caso: mais intenso nos crioulos, menos intenso no português popular do Brasil. Assim, enquanto, nos crioulos portugueses da África, a concordância de gênero com os determinantes, adjetivos e modificadores foi praticamente eliminada, conservando-se apenas, e de forma variável, a marcação do gênero nos nomes dos seres animados que possuem a propriedade de flexão de gênero; em Helvécia, a concordância de gênero entre o nome e os demais constituintes do SN é feita de forma variável, ao passo que a marcação do gênero é categórica nos núcleos nominais flexionáveis em gênero. Essa semelhança fundamental do padrão seguido nos dois processos torna-se, assim, a prova decisiva do caráter crioulizante da variação na concordância de gênero em Helvécia.

Por outro lado, a ideia de que a variação na concordância de gênero que se observa em Helvécia poderia ser o reflexo de uma deriva interna da língua portuguesa deve ser posta de lado, não apenas em função desse significativo paralelo, mas também

em função da visão apresentada na seção 11.2, acima, segundo a qual a morfologia de gênero, ao contrário de se enfraquecer, se fortaleceu ao longo da história interna do português, com o aumento do rendimento funcional do morfema de feminino *-a*. O cabo-verdiano, o Crioulo e o santomense se formaram entre os séculos XV e XVII. Nesse período, a categoria gramatical do gênero estava se expandindo no português. Entretanto, pode-se afirmar que se operou, no processo de constituição desses crioulos, uma verdadeira *déblâcle* na morfologia de gênero e no mecanismo da concordância através do qual essa categoria gramatical se manifesta na sintaxe.

Assim sendo, o que se observa no processo de formação dos crioulos de base portuguesa e no português afro-brasileiro põe em questão o princípio de que a marcação do gênero no português é feita fundamentalmente pelos termos determinantes. Pode-se pensar que assim o é no português contemporâneo, porque a lógica da estrutura interna da língua se sobrepôs à economia básica do processo de prover informação referencial, que constitui a função básica da linguagem humana. Em sua evolução, as línguas naturais desenvolvem todo um aparato estrutural, através de um processo secular que se rege por uma série de fatores de ordem cultural e ideológica, de modo que nem sempre esse processo está em sintonia com a economia da língua enquanto sistema provedor de informação referencial *par excellence*, ou seja, dentro de uma lógica em que a estruturação formal é determinada pelos significados que a forma veicula. Portanto, esta análise revela que a marcação do gênero se faz essencialmente no nome núcleo do SN, não nos termos determinantes. Essa posição só pode ser assumida em função do princípio teórico de que a estruturação da língua é regida a partir da relação linguística essencial entre a forma e o significado (CHOMSKY, 1993, 1995). Sobretudo nos processos em que o *modus linguae* está efetivamente em jogo (como é caso dos processos de transmissão linguística irregular), a relação essencial entre forma e significado se impõe, rompendo as relações que a forma mantém consigo mesma.

No caso do gênero, o processo de erosão gramatical ocorrido durante o processo de transmissão linguística irregular afetou principalmente a concordância de gênero dos determinantes e modificadores com o núcleo nominal do SN, sobretudo nos casos em que essa marcação do gênero veiculava apenas uma classificação gramatical sem nenhuma implicação no plano do significado referencial. A marcação do gênero somente foi preservada, quando possuía um significado referencial — expressando o gênero natural dos indivíduos; mesmo assim, da maneira mais econômica: apenas no núcleo nominal (através da flexão e/ou da composição lexical), eliminando-se ou restringido-se o mecanismo subsidiário (e redundante) da concordância. Fica, assim, explicado por que o progressivo ataque crioulizante sobre os mecanismos sintáticos e mórficos de marcação da categoria gramatical do gênero em português encontrou no núcleo nominal do SN a sua última e mais inexpugnável cidadela.

13 O artigo definido

Alan Baxter

*Norma Lopes**

A variação no uso do artigo definido diante de possessivo é amplamente observada no Brasil, mesmo em contextos urbanos (BAXTER; LOPES, 2004; SILVA, 1996). Por outro lado, na literatura sobre as variedades rurais do português brasileiro, há também a referência ao emprego variável do artigo definido em contextos que não envolvem posse (AMARAL, 1920, p.74; VEADO, 1982, p.37), mas os exemplos são em princípio raros.[1] Porém, entre as várias características crioulizantes do dialeto de Helvécia anotadas por Ferreira (1984), a possibilidade de o sintagma nominal (SN) sem artigo representar referência definida específica chama muito a atenção. Em trabalho de campo nas décadas de 1980 e 1990, essa variável ainda revelou alguma vitalidade, como mostram os seguintes exemplos:

(1) ***Terrero*** *era grande.*

(2) *Pedeu* ***vista****, não é?*

(3) *Io fechô* ***janela****.*

(4) *Puque* ***farinha*** *ai tá ruim!*

(5) *Io num vai pedi* ***caro*** *dele nom.*

Estudos sobre os crioulos de base lexical portuguesa do Atlântico (LUCCHESI, 1993; ALEXANDRE; HAGEMEIJER, 2007; BAPTISTA, 2002, 2007) demonstram que o SN sem artigo desempenha um importante papel na gramática do SN relativamente à referência definida e indefinida. O SN sem artigo é um fenômeno bastante geral nas línguas crioulas, possivelmente um vestígio de reestruturações morfossintáticas nas fases iniciais da história dessas línguas.

---

* O texto é derivado de Baxter e Lopes (2006). Porém, na preparação do presente capítulo, os dados e a sua codificação foram revistos e foram realizadas três novas rodadas. Os resultados apresentados aqui substituem aqueles relatados no referido trabalho anterior.

[1] Também se observa um uso variável do artigo definido no português falado como segunda língua por povos indígenas (AICHENVALD, 2002, p.317; MATTOS E SILVA, 1988, p.106-107).

O surgimento de um sistema referencial que inclui a ausência do artigo definido estaria ligado parcialmente a questões de economia de sistema orientadas pelo grau de previsão de referência atribuído pelos contextos sintático-discursivos e pragmáticos. Na aquisição da segunda língua (doravante L2), a ausência do artigo definido é registrada em vários estudos que envolvem diversas línguas maternas (CAIN; WEBER-OLSEN; SMITH, 1987; KLEIN; PERDUE, 1992) e parece corresponder a uma fase inicial de desenvolvimento da L2. Contudo, a depender das primeiras línguas (doravante L1) envolvidas, a ausência do artigo definido poderia estar relacionada com influências da L1 na estruturação das categorias funcionais da L2 (HAWKINS, 2001, p.239, 244-247) ou, em situações de contato linguístico complexas que envolvem diversas L1 e a aquisição natural de uma L2, poderia estar relacionada simplesmente com a estruturação sistemática da gramática, não sendo necessário apelar para influências da L1 (ROMAINE, 2003, p.419-420).

O presente capítulo apresenta um estudo da variável *artigo definido* no dialeto de Helvécia por meio de dados recolhidos no final da década de 1980 e na primeira metade da década de 1990 (cf. capítulo 5 deste livro).

13.1 Metodologia e hipóteses

Foram analisados dados de onze informantes em quatro faixas etárias:[2] 35 a 45 anos (3 informantes), 60 a 70 anos (3 informantes), mais de 80 anos (4 informantes), e um informante de mais de 100 anos. Um total de 2.655 SNs de referência definida foram depreendidos, registrando-se 80,2% de presença do artigo definido, por meio do pacote GOLDVARB-X. Avaliou-se o efeito de oito variáveis independentes sobre a variável *artigo definido*, que visam a esclarecer a distribuição e a função dessa variável no dialeto e podem ser apresentadas como se segue:

Variável 1, faixa etária do informante: os estudos de outras variáveis morfossintáticas nesse dialeto apontam para a mesma conclusão: a morfossintaxe do padrão urbano, como o emprego das regras de concordância, foi adquirida a partir do início do século XX. A variável *artigo definido* apresentaria o mesmo gênero de configuração?

As variáveis independentes restantes (variáveis 2 a 8), todas linguísticas, estão motivadas pelas discussões em estudos sobre o SN e o substantivo sem artigo, no crioulo de Cabo Verde (LUCCHESI, 1993; BAPTISTA, 2002, 2007) e no

[2] O *corpus* analisado inclui principalmente dados recolhidos em 1994 mas, para a configuração da faixa 4, e para incluir dados de um informante especial (filho de escravos), foram incluídos também dados de recolhas de 1987-1988: dois informantes da mesma geração dos informantes da faixa 4 de 1994 e dados de M.L., filho de escravos.

crioulo de São Tomé (ALEXANDRE; HAGEMEIJER, 2007; LUCCHESI, 1993). As tipologias dessas línguas com substratos das famílias linguísticas africanas atlântico-ocidentais (subgrupo do níger-congo), no crioulo de Cabo Verde (doravante CCV), e benue-congo e banto, no crioulo de São Tomé (doravante CST), podem ser de relevância para a análise do dialeto de Helvécia que tem substratos benue-congo e banto. No passado da comunidade de Helvécia, houve uma fase em que variedades de português L2 faladas por africanos serviram de dados linguísticos primários para a aquisição do português L1, processo que teria facilitado a incorporação de características crioulas (cf. seção 2.2.2 do capítulo 2).

A **variável 2**, **animacidade do substantivo**, é proposta porque no CST a interpretação semântica do SN sem artigo manifesta alguma sensibilidade ao traço [+humano] e pode, sobretudo em posição de sujeito, levar à interpretação 'SN definido singular' (ALEXANDRE; HAGEMEIJER, 2007, p.49). Em CCV, detecta-se uma relação entre, por um lado, os três fatores animacidade, definitude e flexão de plural e, por outro lado, a ausência do artigo definido (BAPTISTA 2007, p.86). Também, seguindo o pensamento de Lyons (1999) de que os referentes humanos definidos são mais salientes na percepção humana, levanta-se a hipótese de que a animacidade poderia favorecer a ausência do artigo definido.

A **variável 3**, **contabilidade do substantivo**, visa a determinar se a noção de contabilidade, que diz respeito ao número, em contraste com a noção de massa, poderia exercer uma influência sobre as variantes do artigo definido. A análise é inspirada parcialmente na discussão feita por Baptista (2007, p.73-74) a respeito das implicações do parâmetro de mapeamento nominal proposto por Chierchia (1998), que atribui ao substantivo contável sem artigo uma denotação de "massa" no sentido de "classificação de tipo" ou "classe". Baptista observa que, no CCV, o substantivo contável sem artigo é tratado como uma entidade individual.

A **variável 4**, **referência do SN (única ou específica)**, avalia o condicionamento sobre a variável dependente exercido pela referência única. O fator pragmático da referência única facilita o uso do substantivo sem artigo no CST (ALEXANDRE; HAGEMEIJER, 2007, p.53-54) e no CCV (BAPTISTA 2007, p.74).

A **variável 5**, **número do SN (singular ou plural)**, visa a avaliar o efeito do número do SN sobre a variável dependente. No CST, o substantivo sem determinante é tipicamente singular (ALEXANDRE; HAGEMEIJER, 2007, p.52), sendo o substantivo [+humano] especialmente comum nesses casos (p.49). No CCV, Baptista (2002, p.41; 2007, p.84, 86) observa que a flexão plural no SN está ligada à definitude e à animacidade.

A **variável 6**, **presença de outros modificadores no SN**, avalia se a presença de outros modificadores no SN, capazes de atribuir referencialidade, inibe a presença do artigo pleno. Esse efeito é registrado no CCV (LUCCHESI, 1993, p.90; BAPTISTA, 2007, p.81).

A **variável 7**, **determinante no SN anafórico prévio**, avalia o grau de condicionamento exercido pelos determinantes de SNs anafóricos antecedentes. Relativamente ao CCV, Lucchesi (1993, p.81) e Baptista (2007, p.75-76) observaram que a referência anafórica a um antecedente introduzido por meio do artigo definido podia ser retomada ou pelo artigo explícito ou pela sua ausência. Segundo Baptista, a presença prévia do artigo indefinido é desnecessária, pois no CCV o substantivo pode aparecer sem artigo mesmo na primeira ocorrência.

A **variável 8**, **função sintática do SN**, avalia o condicionamento da variável dependente por aspectos estruturais e funcionais no nível da frase. No CCV (BAPTISTA, 2007, p.77, 81, 86), o SN singular sem artigo é interpretado preferencialmente como definido, quando ocorre em posição de sujeito, e como indefinido, quando ocorre em posição de objeto. Em contrapartida, no CST, Alexandre e Hagemeijer (2007, p.55) não encontraram essa assimetria na distribuição dos SNs sem artigo nos argumentos do verbo.

13.2 Análise dos dados

Na análise global do *corpus* de Helvécia, só as variáveis *animacidade do substantivo* e *contabilidade do substantivo* foram rejeitadas pelo programa GOLDVARB por serem estatisticamente não significativas. A seguir, são analisados os resultados quantitativos das variáveis independentes selecionadas como estatisticamente relevantes.

13.2.1 A faixa etária: o perfil diacrônico da variável *artigo definido*

A distribuição geracional da variável *artigo definido* revela um nítido perfil aquisicional. A Tabela 1 apresenta os resultados das três faixas plenas, ao lado da “faixa 5”, representada pelo informante M.L., filho de escravos.[3, 4]

Tabela 1 - Efeito da faixa etária sobre o uso do artigo definido em Helvécia-BA

FAIXA ETÁRIA	Nº. DE DADOS	%	P.R.
Faixa 2 (35 a 45 anos)	1090/1147	95%	.750
Faixa 3 (60 a 70 anos)	514/657	78.2%	.479
Faixa 4 (mais de 80 anos)	423/666	63.5%	.212
“Faixa 5” (mais de 100 anos)	102/185	55.1%	.143

Log likelihood = -851.435 Significância = .000 [Rodada 1]

[3] O informante M. L. estava, possivelmente, com 105 anos, em 1987.

[4] As faixas etárias estão enumeradas de 2 a 5 porque os dados provêm de *corpus* maior em que há cinco faixas. Neste estudo, não se considerou a faixa 1.

Nas faixas 4 e 5, a gramática do artigo definido está numa fase inicial, em que a ausência do artigo definido é significativa. Em contrapartida, nas faixas 2 e 3, há um aumento considerável do artigo pleno, reflexo de uma gradual aquisição de formas mais próximas do padrão. Consideramos que a presença do substantivo sem artigo é um vestígio de um sistema prévio em que prevaleceu um reflexo das fases iniciais da comunidade, quando dados de português L2 *interlíngua* foram produtivos (BAXTER; LUCCHESI, 1999).

13.2.2 Variáveis linguísticas

As variáveis linguísticas independentes revelam vários fatos curiosos muito significativos a respeito da distribuição e das funções da variável *artigo definido* no dialeto de Helvécia. Por um lado, os perfis de determinadas variáveis independentes relembram as tendências referidas nos crioulos de base portuguesa do Atlântico. Por outro lado, outras variáveis apontam para características que diferenciam o dialeto de Helvécia das tendências registradas nessas línguas.

13.2.2.1 Aspectos semânticos do substantivo

A seleção da variável *referência do SN* como estatisticamente significativa, cujos resultados estão apresentados na Tabela 2, confirma a hipótese de que o conhecimento compartilhado pelos interlocutores facilita bastante o uso artigo definido nos SNs que têm referentes claramente identificáveis, como ocorre com o substantivo *lua*, no exemplo seguinte:

(6) *tava oiano na lua assim ói, **lua** tá traz do sol uma bandinha*

Tabela 2 - Efeito da referência do SN sobre o uso do artigo definido em Helvécia-BA

REFERÊNCIA DO SN	Nº. DE DADOS	%	P.R.
referência [+único]	89/130	68.5%	.218
referência [-único]	2040/2524	80.8%	.516

Log likelihood = -851.435 Significância = .000 [Rodada 1]

O percentual de 68,5%, com peso relativo de .218, indica que a referência única desfavorece o uso do artigo definido. Registra-se, assim, um paralelo com os fatos do CCV e do CST referidos na seção 13.1 acima. Por outro lado, a referência não única tem um efeito neutro sobre a variável dependente.

13.2.2.2 O número do SN

O SN plural, exemplificado em (7) abaixo, revela-se um fator bastante favorável ao emprego do artigo definido, com um peso relativo de .800 (cf. Tabela 3).

Ao contrário, o SN singular desfavorece ligeiramente o uso do artigo definido (p.r. 458).[5]

(7) *num contano* ***as mulhê*** *não*

Tabela 3 - Efeito do número do SN sobre o uso do artigo definido em Helvécia-BA

NÚMERO DO SN	N^{O.} DE DADOS	%	P.R.
SN plural	266/286	93%	.800
SN singular	1863/2369	78.6%	.458

Log likelihood = -851.435 Significância = .000 [Rodada 1]

Esse panorama relembra as pesquisas sobre o CST (ALEXANDRE; HAGEMEIJER, 2007) que chamam a atenção para um vínculo formal entre as noções semânticas de [+definitude] e [+plural], que nesse crioulo estão encapsuladas no determinante definido plural *inen*. Ao mesmo tempo, à semelhança do CST (ALEXANDRE; HAGEMEIJER, 2007) e do CCV (BAPTISTA, 2007), o resultado aponta para uma conexão entre o SN definido singular e ausência do artigo definido.

Numa tentativa de avaliar o potencial efeito da variável *número do SN* nas diferentes gerações, essa variável foi cruzada com a *faixa etária*. Os resultados estão apresentados na Tabela 4. De uma perspectiva diacrônica, é evidente que o SN plural lidera o incremento do uso do artigo definido. Em todas as faixas, o SN plural favorece o uso do artigo definido; um efeito que cresce em força, conforme diminui a idade: "faixa 5" (p.r. .535), faixa 4 (p.r. .632), faixa 3 (p.r. .650) e faixa 2 (p.r. 961). Em contrapartida, o SN singular só se torna favorável ao artigo definido na faixa 2.

Tabela 4 - Efeito do número do SN sobre o uso do artigo definido por faixa etária em Helvécia-BA

FAIXA ETÁRIA	FATOR	N^O DE DADOS	%	P.R.
Faixa 2 (35-45 anos)	SN plural	126/127	99.2%	.961
	SN singular	964/1020	94.5%	.708
Faixa 3 (60-70 anos)	SN plural	79/90	87.8%	.650
	SN singular	435/567	76.7%	.449
Faixa 4 (80 + anos)	SN plural	52/59	88.1%	632
	SN singular	371/607	61.1%	.177
"Faixa 5" (100 + anos)	SN plural	9/10	90%	.535
	SN singular	93/175	53.1%	.119

Log likelihood= -849.004 Significância= .000 [Rodada 2]

[5] Este resultado corrige valores apresentados em Baxter e Lopes (2006, p.60) em que se atribuiu um valor ainda mais inibidor ao SN singular (p.r. .22). Os resultados do estudo atual aproximam-se daqueles do estudo piloto dos dados urbanos de Salvador (BAXTER; LOPES, 2004): SN plural, p.r. .71; SN singular, p.r. .45.

Também foram cruzadas as variáveis independentes *número do SN* e *animacidade* (cf. Tabela 5), numa tentativa de avaliar os possíveis vínculos entre o artigo definido nulo e o SN singular com núcleo [+humano], por um lado, e entre o artigo definido pleno e o SN plural com núcleo [+humano], por outro (ALEXANDRE; HAGEMEIJER, 2007, p.41, 49, 52, para o CST).

Tabela 5 - Efeito do número e da animacidade do SN sobre o uso do artigo definido em Helvécia-BA

SUBSTANTIVO	Nº. DE DADOS	%	P.R.
+humano, plural	150/160	93.8%	.801
-humano, plural	116/125	92.8%	.816
+humano, singular	333/527	63.2%	.429
-humano, singular	1518/1827	83.1%	.465

Log likelihood = -849.333 Significância = .000 [Rodada 3]

Embora o SN singular com núcleo [+humano] singular seja o fator mais favorável à ausência do artigo, a diferença de pesos relativos entre esse fator e o SN singular com núcleo [–humano] é marginal. No total, embora o SN [+humano, +singular] do dialeto de Helvécia manifeste uma vaga semelhança com os fatos do CST, o número plural continua a ser o fator predominante favorecedor do artigo definido pleno, enquanto o SN singular prefere o artigo definido nulo.

13.2.2.3 Presença de outro constituinte capaz de induzir a referência definida

Na análise do efeito da presença de outros modificadores, foram considerados os sete fatores listados e exemplificados a seguir:

(8) Nenhum outro modificador: *ficô eu pra contá* ***o caso***

(9) Oração relativa: ***a casa que eu morava*** *é lá*

(10) Adjetivo pré-nominal: *é* ***a merma coisa***

(11) Adjetivo pós-nominal: ***a terra vermelha****, ela aguenta*

(12) Sintagma preposicional: *eu vi* ***distância do caminho***

(13) Possessivo: ***o meu pai*** *ajuntô cum ela*

(14) Advérbio locativo: ***a bica ali*** *secô*

Na rodada geral, esta variável foi selecionada como significativa, porém os resultados apresentados na Tabela 6 sugerem que esses fatores não se comportam da mesma maneira.

Tabela 6 - Efeito da presença de outros modificadores no SN sobre o uso do artigo definido em Helvécia-BA

OUTROS MODIFICADORES NO SN	Nº. DE DADOS	%	P.R.
Nenhum outro	1400/1640	85.4%	.561
Oração relativa	70/84	83.3%	.600
Adjetivo pré-nominal	27/29	93.1%	.801
Adjetivo pós-nominal	53/60	88.3%	.633
Sintagma preposicional	308/401	76.8%	.446
Possessivo	79/191	41.4%	.131
Advérbio locativo	150/169	88.8%	.533

Log likelihood = -851.435 Significância = .000 [Rodada 1]

Detectam-se aí duas tendências. Por um lado, há algumas evidências da inibição do uso do artigo definido na presença de outros elementos atribuidores de referência, efeito notado por Lucchesi (1993) no CCV. Isso é notável com o possessivo, um forte portador de definitude, com p.r. .131.[6] Esse mesmo efeito é marginalmente manifesto com o sintagma preposicional (p.r. .446), que inclui também estruturas possessivas preposicionais, como no exemplo (5).

Em realidade, os possessivos pré-nominais têm um comportamento muito diferente daquele dos outros modificadores contemplados aqui e são capazes de atribuir um grau de referência definida suficiente para dispensar a presença do artigo definido, uma característica geral no português brasileiro.[7] Já a presença de outro constituinte atribuidor de referência realmente não inibe o uso do artigo definido, como se observa com o adjetivo pré-nominal (p.r. .801), a oração relativa (p.r. .600) e o adjetivo pós-nominal (p.r. .633).

O comportamento dos advérbios locativos, que levam a uma redução do uso do artigo relativamente maior que os outros constituintes (p.r. .533), pode ser bem analisado de uma perspectiva aquisicional. No dialeto de Helvécia, há evidências de que o locativo atua como um demonstrativo, capaz de funcionar como artigo (BAXTER; LUCCHESI; GUIMARÃES, 1997, p.13-14). Seguindo a proposta de Cerqueira (1993) de que os locativos funcionam como demonstrativos nas primeiras fases de desenvolvimento do SN na aquisição L1 do português brasileiro, os dados apresentados aqui representariam uma fase de transição em que o artigo definido está sendo desenvolvido em estruturas que previamente dependiam do locativo

---

[6] Baxter e Lopes (2004) registraram um resultado semelhante (p.r. .18) em dados urbanos de Salvador.

[7] Em Baxter, Lucchesi e Guimarães (1997, p.24-26), o possessivo foi analisado como um argumento interno do NP, gerado como especificador. Contudo, a capacidade de atribuição de referência definida neste *corpus* vai ao encontro da perspectiva de Castro e Costa (2003), que consideram que o possessivo no português europeu apresenta qualidades de núcleo (*head*).

pós-nominal como marcador das propriedades referenciais do substantivo. Nessa fase transicional, esse papel pretérito do locativo começa a recuar enquanto se desenvolve a estrutura pré-nominal.[8]

Como o artigo definido é altamente favorecido nos SNs plurais, detecta-se na coocorrência do artigo definido e modificadores pós-nominais nos dados de Helvécia um paralelo com o estudo de Alexandre e Hagemeijer (2007, p.41) sobre o CST, no qual observaram que um grande número de substantivos estavam modificados simultaneamente pelo marcador plural definido *inen* e por modificadores pós-nominais que atribuíam especificidade como, por exemplo, a oração relativa, o adjetivo e o possessivo.

Sob a perspectiva do desenvolvimento diacrônico do dialeto, parece razoável propor que os modificadores favorecedores da presença do artigo definido constituem contextos estruturais em que o uso do artigo foi primeiramente adquirido.[9]

13.2.2.4 Interação com o determinante do SN correferencial prévio

Na avaliação do efeito dessa variável, foram considerados quatro fatores, exemplificados de (15) a (18), cujos resultados estão na Tabela 7:

(15) Artigo definido: *quando eu tava com* ***o machado****, quando eu bati* ***o machado*** *no pau, quando bateu fez <taco>!*

(16) Sem artigo definido: *Essas menina é... é pa barrê* ***terreiro****. [...] poque* ***terrêro*** *era grande*

(17) Demonstrativo: *Qu'ocê num contá* ***esse garucho****. Pá* ***garucho*** *levá*!

(18) Artigo indefinido: *Fica* ***uma parte*** *pra ali,* ***ôtra parte*** *pra cá!*

Tabela 7 - Efeito do determinante no SN anafórico prévio sobre o uso do artigo definido em Helvécia-BA

DETERMINANTE: SN PRÉVIO	Nº. DE DADOS	%	P.R.
Artigo definido pleno	411/462	89.0%	.610
Artigo definido	104/204	51.0%	.288
Demonstrativo	27/37	73.0%	.345
Artigo indefinido	37/43	86.0%	.504

Log likelihood = -851.435 Significância = .000 [Rodada 1]

[8] SCHWEGLER (2007, p.21) aponta para um mecanismo semelhante no palenquero (crioulo de base lexical espanhola, falado na região de Cartagena, na Colômbia).

[9] Esta ideia é parcialmente confirmada em resultados de análises independentes de cada faixa etária. Assim, por exemplo, numa análise independente da faixa etária 4 (Log liklihood = -285.372; significância = .000), o uso do artigo definido em SNs com orações relativas registra um peso relativo quase neutro (p.r..48). Em contrapartida, em análises independentes das faixas etárias 3 e 2, o uso do artigo definido é categórico nesse contexto.

Os resultados revelam uma interação complexa entre o determinante do SN correferente prévio (SN1) e o determinante do SN correferente subsequente (SN2). Quando o SN1 leva o artigo definido pleno ou ∅, há um efeito sequencial: é favorecida a repetição desse determinante no SN2.[10] Em contrapartida, a presença do demonstrativo no SN1 favorece a ausência do artigo definido no SN2. Como o demonstrativo atribui os dois valores referenciais de dêixis e definitude, a referência do SN está suficientemente fundamentada para dispensar o determinante na ocorrência subsequente desse SN.

Por outro lado, o artigo indefinido no SN1 exerce um efeito quase neutro sobre a presença do artigo pleno no SN2. Nesse caso, há um paralelo com os fatos do CCV (LUCCHESI, 1993; BAPTISTA, 2002, 2007), e uma diferença em relação ao CST, que requer um passo intermediário no qual o SN2 aparece com o marcador de especificidade *se*, e só na ocorrência posterior aparece o SN sem determinante definido (ALEXANDRE; HAGEMEIJER, 2007, p.55).

13.2.2.5 Função sintática do SN

Com esta variável, foram investigados os efeitos de sete fatores listados nos exemplos (19) a (25) abaixo,[11] cujos resultados figuram na Tabela 8.

(19) Sujeito: ***mãezinha de nós*** *caregava nós assim*

(20) Objeto direto: *eu vi* ***distância do caminho***

(21) Predicativo: *Ma lá é* ***casa dela!***

(22) Adjunto adverbial: *Agora só quand'ele vai lá* ***na venda***

(23) Adjunto adnominal: *non dou dinheiro* ***de dizmo***

(24) Objeto indireto: [eles] *vendeu* ***pas firma****, vendeu* ***pa fazendero***

(25) SN independente: *Eu......* ***meu irmão****...*[12]

A análise revela que o SN sem artigo definido aparece numa ampla gama de funções, mas nota-se que o artigo definido pleno só é consistentemente favorecido na função de adjunto adnominal, enquanto o SN independente nem o favorece nem o desfavorece. Em contrapartida, todos os fatores restantes desfavorecem o

[10] O fenômeno relembra o processamento paralelo detetado na variável presença da flexão de plural no SN, observado por Scherre (1988).

[11] As funções de tópico e aposto não foram tratadas nesta análise, tendo em vista o baixo número de dados.

[12] Resposta à pergunta do entrevistador: "Quem faz esse trabalho?".

uso do artigo definido, e as funções de sujeito e objeto direto são os menos inibidores, com valores próximos.[13]

Tabela 8 - Efeito da função sintática do SN sobre o uso do artigo definido em Helvécia-BA

FUNÇÃO SINTÁTICA DO SN	Nº. DE DADOS	%	P.R.
Sujeito	492/690	71.3%	.456
Objeto direto	402/531	75.7%	.445
Predicativo	78/119	65.5%	.335
Adjunto adverbial	827/871	94.9%	.683
Adjunto adnominal	169/230	73.5%	.208
Objeto indireto[14]	57/68	83.8%	.341
Frase independente	47/63	74.6%	.494

Log likelihood = -851.435 Significância = .000 [Rodada 1]

Quanto ao comportamento favorecedor do adjunto adverbial, há duas explicações plausíveis. Em primeiro lugar, o artigo pleno teria de aparecer porque o adjunto adverbial não está governado pelo verbo por meio de *c-command* e, por isso, a predicação não pode atribuir referencialidade ao SN. Estendendo essa linha de argumentação, muitos adjuntos adverbiais são SNs preposicionados; portanto, a presença da preposição pode motivar o peso favorável do adjunto adverbial. Porém o SPrep adjunto adnominal, governado por um substantivo, recebe desse uma referência suficiente para dispensar o artigo pleno. Ao mesmo tempo, a distribuição de determinadas preposições pode ser relevante. Os possessivos pós-nominais figuram entre os adjuntos adnominais, e os possessivos levam uma forte carga referencial. O possessivo pós-nominal poderia contribuir para o peso desfavorável do adjunto adnominal.

O perfil geral dos resultados aponta para o desenvolvimento inicial do artigo definido pleno nos SNs que são periféricos ao SV ou independentes. Ao mesmo tempo, os resultados registrados pelas funções de sujeito e objeto direto parecem ser superficialmente semelhantes aos dados do SN no CST, que não apresenta assimetria na distribuição dos SNs sem determinante de referência definida que figuram como argumentos internos do verbo (ALEXANDRE; HAGEMEIJER, 2007).

[13] Baxter e Lopes (2004) obtiveram resultados parcialmente semelhantes a esses na análise de dados de Salvador: adjunto adverbial favorece o artigo definido pleno (p.r. .60), enquanto são desfavoráveis o sujeito (p.r. .45), o adjunto adnominal (p.r. .29) e o predicativo (p.r. .16), porém há várias diferenças. Por um lado, o objeto indireto é o fator mais favorecedor (p.r. .74), seguido do objeto direto (p.r. .60). Por outro lado, o SN independente é altamente desfavorável à presença do artigo definido pleno. Consideramos que essas diferenças refletem as diferentes fases de desenvolvimento das duas variedades linguísticas em questão.

[14] O objeto indireto neste dialeto é também representado por um SN não preposicionado, na configuração VERBO+OI+OD (cf. capítulo 18 deste livro).

13.3 Conclusão

O presente capítulo atribui mais um dado ao conjunto de variáveis que testemunham processos de reestruturação morfossintática no passado do dialeto de Helvécia, provocados pela natureza dos dados linguísticos primários disponíveis para a aquisição do português L2 e L1 nas primeiras décadas da fundação da Colônia Leopoldina, quando a maioria dos escravos era de procedência africana. A variável apresenta um perfil diacrônico aquisicional em relação ao artigo definido pleno, e há evidências de uma fase pretérita no dialeto, quando o SN definido sem artigo teria sido mais frequente. Ao mesmo tempo, a análise apontou para algumas semelhanças com a tipologia do SN sem artigo definido no CVS e no CST, e também detectou algumas características idiossincráticas do dialeto em foco. No *corpus* analisado, a variável está amplamente distribuída num SN que ainda está em desenvolvimento sob influência de variedades do português brasileiro mais próximas da norma padrão.

14 A concordância verbal

Dante Lucchesi

Alan Baxter

*Jorge Augusto Alves da Silva**

A concordância verbal de número e pessoa ao lado da concordância nominal de número são os aspectos da gramática da língua portuguesa no Brasil que exibem os mais amplos processos de variação, distinguindo qualitativamente a variante brasileira do português da variante europeia, onde tal variação é, na melhor das hipóteses, residual e periférica. Para além de ser um dos pontos mais focalizados pelas análises sociolinguísticas, a concordância verbal está no centro dos debates acerca da relevância do contato entre línguas na formação da realidade linguística brasileira.

No que concerne ao paradigma da flexão verbal, o português brasileiro culto se diferencia da tradição normativa da seguinte maneira: enquanto nesta, cada pessoa do discurso tem uma forma verbal morficamente diferenciada, no padrão verbal consagrado pelo uso culto brasileiro, reduz-se a quatro formas, como esquematizado no Quadro 1.

NORMA PADRÃO	NORMA CULTA BRASILEIRA
eu trabalh**o**	eu trabalh**o**
tu trabalha**s**	você trabalha
ele/ela trabalha	ele/ela trabalha
nós trabalha**mos**	nós trabalha**mos**
vós trabalha**is**	vocês trabalha**m**
eles/elas trabalha**m**	eles/elas trabalha**m**

Quadro 1 - O paradigma da flexão verbal no padrão normativo e no português brasileiro culto

* A autoria deste capítulo é dividida da seguinte maneira: a introdução ficou a cargo de Dante Lucchesi; a seção 14.1. é de autoria de Dante Lucchesi e Alan Baxter; a seção 14.2. é de Jorge Augusto Alves da Silva e Dante Lucchesi; e seção 14.3. é de autoria de Dante Lucchesi.

Considerando as categorias gramaticais de pessoa e número no paradigma do uso culto brasileiro, pode-se afirmar que: a 1ª pessoa do singular é marcada na forma verbal quanto à categoria de pessoa; a 2ª e a 3ª pessoas do singular se relacionam com formas verbais não marcadas morficamente; as formas verbais da 2ª e da 3ª pessoas do plural são marcadas apenas quanto ao número; e a forma verbal da 1ª pessoa do plural é marcada cumulativamente nas categorias gramaticais de pessoa e número, como se pode ver no Quadro 2.

PESSOA DO DISCURSO	TRAÇOS DE PESSOA E NÚMERO	FORMA VERBAL
1ª pessoa do singular	[+pessoa / -plural]	eu trabalh**o**
2ª pessoa do singular	[-pessoa / -plural]	você trabalha
3ª pessoa do singular	[-pessoa / -plural]	ele/ela trabalha
1ª pessoa do plural	[+pessoa / +plural]	nós trabalha**mos**
2ª pessoa do plural	[-pessoa / +plural]	vocês trabalha**m**
3ª pessoa do plural	[-pessoa / +plural]	eles trabalha**m**

Quadro 2 - Os traços de pessoa e número no paradigma da flexão verbal no português brasileiro culto

O morfema de pessoa *-o* e seu alomorfe *-i* ocorrem apenas no presente, no pretérito perfeito e no futuro do presente do indicativo; nos demais tempos, essa marcação mórfica se neutraliza, não restando qualquer distinção na forma verbal que se refere a todas as pessoas do discurso no singular. No que concerne à 2ª pessoa do singular, o pronome *tu* conjugado com uma forma marcada pelo morfema *-s* ainda subsiste em algumas poucas regiões do Brasil, principalmente no Sul, sofrendo uma crescente concorrência da estrutura com *você*, ou de uma estrutura em que *tu* se combina com uma forma morficamente não marcada. Porém, na maior parte do território brasileiro, predomina a estrutura de *você* com a forma verbal não marcada e, na norma popular e rural de muitas regiões, *tu* se combina categoricamente com a forma verbal não marcada morficamente para pessoa e número.

O pronome *nós* vem progressivamente sendo substituído, em função do processo de gramaticalização da expressão nominal *a gente* (cf. capítulo 19 deste livro). Tal processo reduz ainda mais a flexão de número e pessoa do verbo, na medida em que a expressão nominal *a gente*, gramaticalizada como forma variante do pronome de 1ª pessoa do plural, se combina com as formas verbais não marcadas da 3ª pessoa do singular. Esse processo de mudança em curso atinge tanto a norma culta quanto a norma popular, o que tem reduzido bastante o uso da forma marcada em *-mos*. Entretanto, na norma popular, o quadro de variação é mais complexo, porque, ao lado dessas duas estruturas (*nós trabalhamos* e *a gente trabalha*), se observam tanto a combinação de *a gente* com a forma flexionada em *-mos* (*a gente*

trabalhamos), quanto o uso de *nós* com a forma verbal não marcada morficamente (*nós trabalha*).

O morfema *-m* de número é aquele que sofre o processo mais amplo de variação no português brasileiro (PB). Segundo Scherre e Naro (1997, p.93), a concordância verbal e nominal, no que concerne à categoria gramatical de número, é geral e amplamente documentada no PB. Esses autores a definem como um caso de variação inerente, com base no conceito proposto por Labov (1972a, p.183-259). Esse processo de variação, que atinge a regra de concordância verbal com a 2ª e a 3ª pessoas do plural, atinge tanto a norma culta quanto a norma popular, porém há que se destacar a enorme diferença entre os dois polos desse contínuo relativamente à frequência de uso do morfema de número na norma culta, que é superior aos noventa pontos percentuais, enquanto, no português afro-brasileiro, fica abaixo dos vinte pontos percentuais (ver 14.2 adiante). O Quadro 3 apresenta esquematicamente os processos de variação que ocorrem no paradigma da flexão verbal nas duas grandes normas do português brasileiro.

NORMA CULTA BRASILEIRA	NORMA POPULAR BRASILEIRA
eu trabalh**o**	eu trabalh**o**
você trabalha (~ tu trabalha**s** ~ tu trabalha)	você trabalha ~ tu trabalha
ele/ela trabalha	ele/ela trabalha
nós trabalha**mos** ~ a gente trabalha	nós trabalha ~ nós trabalha**mo(s)** a gente trabalha ~ a gente trabalha**mo(s)**[1]
vocês trabalha**m** (~ vocês trabalha)	vocês trabalha ~ vocês trabalha**m**
eles/elas trabalha**m** (~ eles/elas trabalha)[2]	eles/elas trabalha ~ eles/elas trabalha**m**

Quadro 3 - Processos de variação no paradigma da flexão verbal no português brasileiro culto e no português popular do Brasil

O *continuum* que se verifica em relação à frequência de uso do morfema de número, desde a norma urbana culta, de um lado, até as comunidades rurais afro-brasileiras, no extremo oposto, passando pelo que se pode chamar de norma urbana semiculta e variedades populares urbanas e rurais do PB, constitui uma significativa evidência empírica de como o contato entre línguas afetou as diversas variedades do PB. As variedades populares que têm a sua formação mais fortemente marcada pelo contato entre línguas são aquelas que apresentam as menores frequências de uso do morfema de número, ao passo que as variedades urbanas culta e semiculta,

[1] É muito comum a supressão do *-s* final do morfema *-mos* no português popular do Brasil.

[2] Vale ressaltar que, para a norma culta, considerada a análise feita na amostra do NURC do Rio de Janeiro (GRACIOSA, 1991), o nível de variação na aplicação da regra de concordância com a 3ª pessoa do plural é inferior a 5%. Esse nível deve se manter em relação à 2ª pessoa do plural.

que só indiretamente foram afetadas pelo contato em sua formação histórica, são aquelas que apresentam os maiores índices de aplicação da regra de concordância. Além disso, na norma popular, as diversas comunidades de fala apresentam um cenário de mudança em curso no sentido do incremento da aplicação da regra de concordância, o que vai de encontro à hipótese lançada por Naro (1981) de que no PB estaria em curso um processo de perda das marcas de concordância e desautoriza a visão da ação de uma deriva secular, defendida por Naro e Scherre (1993, 2007).

Portanto, o estudo da concordância verbal nas comunidades rurais afro-brasileiras isoladas revela-se um campo bastante profícuo para se estabelecerem relações empiricamente motivadas do contato entre línguas que marca a história sociolinguística do Brasil com o quadro de variação morfossintática que se observa atualmente. Nesse sentido, um dos fatos mais significativos é a variação na concordância verbal relacionada com a 1ª pessoa do singular, que será objeto da seção 14.1 deste capítulo. Esse processo de variação se circunscreve, no conjunto das comunidades rurais afro-brasileiras isoladas aqui analisadas, à comunidade de Helvécia. Esse fato sinaliza para a hipótese de que na formação do PB, pelo menos em algumas localidades em que o contato entre línguas foi mais intenso, a erosão gramatical atingiu todo o paradigma de flexão de pessoa e número do verbo.

A seção 14.2 trata do uso do morfema de número na 3ª pessoa do plural e revela um cenário de reduzidíssima aplicação da regra de concordância verbal entre os falantes mais velhos das comunidades rurais afro-brasileiras isoladas; quadro esse que vai se esbatendo com o crescente uso do morfema de número entre os falantes mais jovens. Fica evidente o contraste entre esse cenário claro de mudança no sentido do aumento do emprego da regra de concordância com o que se depreende dos padrões urbanos de comportamento linguístico, para os quais o prognóstico mais plausível é o de variação estável (SCHERRE; NARO, 1997), sob o qual se podem identificar influxos contrários (NARO; SCHERRE, 1991). Essas diferenças, tanto na frequência de emprego das variantes, quanto nas tendências de mudança em curso, ratificam a visão de uma polarização sociolinguística do Brasil, como sistematizado por Lucchesi (1994, 1998, 2001a, 2002a, 2002b, 2006).

A análise da variação da concordância verbal com a 1ª pessoa do plural, apresentada na seção 14.3, também traz importantes evidências a favor da relevância do contato entre línguas na formação histórica do PB, bem como da polarização sociolinguística do Brasil. Nela se demonstrará que, diferentemente do que ocorreu na norma culta, a redução no uso do morfema cumulativo *-mos* nas comunidades afro-brasileiras isoladas não decorre da substituição do pronome *nós* pela expressão *a gente*, já que os falantes mais velhos, que ainda conservam mais o pronome *nós*, são exatamente aqueles que aplicam menos a regra de concordância verbal. Resta, então, como válida apenas a hipótese de que o quadro atual de variação na aplicação da regra de concordância verbal no português brasileiro nada mais é do que o

resultado da redução no uso da morfologia flexional associada às regras de concordância; redução esta que constitui um processo virtualmente universal nas situações de contato maciço entre línguas (cf. capítulo 3 deste livro).

Portanto, o panorama da variação na concordância verbal nas comunidades rurais afro-brasileiras isoladas que aqui se delineia tem como pano de fundo um amplo processo de variação que constitui uma das mais significativas diferenças entre o português do Brasil e o português europeu, na medida em que, para o segundo, diferentemente do que sugerem Naro e Scherre (2007), não há qualquer evidência efetiva da existência de um quadro de variação que sequer se aproxime do quadro brasileiro. Ao contrário, descortina-se um cenário de ampla vitalidade no uso dos morfemas de pessoa e número do verbo, com particular destaque para o morfema de 2ª pessoa do singular, que tem um valor social notável, como marca de intimidade; para a vitalidade do uso do morfema de 2ª pessoa do plural no interior da região Norte de Portugal; e para a concordância semântica que tem disseminado o uso da estrutura *a gente trabalhamos* mesmo entre os segmentos escolarizados da capital, Lisboa.

Ao invés de um processo de perda progressiva da morfologia verbal de pessoa e número, fruto de uma deriva secular, a caracterização sociolinguística do Brasil impõe um cenário muito mais complexo, do qual emerge a polarização entre a maioria da população, em cuja fala se vêm atenuando os reflexos de um pretérito processo de erosão gramatical resultante do contato entre línguas, frente aos padrões de fala das classes superiores, em que os influxos da interação com as classes baixas convive com forças estruturais de vária ordem e com a pressão social da normatização linguística.

Nas análises que se apresentam a seguir sobre a variação na concordância verbal nas pessoas do discurso que se conjugam com um morfema flexional específico no português brasileiro, tendo como universo de observação as comunidades rurais afro-brasileiras do interior do Estado da Bahia, parte-se da hipótese de que, no processo de formação dessas comunidades de fala, a morfologia flexional de pessoa e número do verbo existente na língua-alvo sofreu um violento processo de erosão, na proporção do grau de transmissão linguística irregular presente na gênese de cada comunidade. Tal processo de erosão chegou a atingir a 1ª pessoa do singular nos casos de contato mais radical, como no caso da comunidade de Helvécia, para a qual se recolheram indícios de um passado crioulo (cf. capítulo 2 deste livro). Porém, diferentemente do que aconteceu, por exemplo, nos crioulos africanos de base portuguesa, essa erosão não foi de monta a eliminar totalmente a flexão de pessoa e número do verbo, mas produziu um amplo e profundo quadro de variação qualitativamente distinto de qualquer processo de variação residual e não estruturada que eventualmente possa ser identificado em qualquer momento da história das variedades cultas e populares do português europeu.

Tal processo de variação ainda se reflete fortemente no português popular brasileiro, sobretudo em suas variedades mais isoladas do interior do país. Entretanto, tem se verificado, nas últimas décadas, uma tendência à implementação do uso de morfemas flexionais do verbo em função da influência linguística e cultural dos grandes centros urbanos sobre todas as regiões do país, graças sobretudo ao enorme poder dos grandes meios de comunicação de massa, mas também ao deslocamento populacional e à expansão do sistema de educação pública (cf. capítulo 1 deste livro). Portanto, em todas as análises sociolinguísticas realizadas nas comunidades afro-brasileiras isoladas, revelou-se um processo de mudança em curso em favor da implementação das regras de concordância verbal.

14.1 A concordância verbal com a primeira pessoa do singular

Nesta seção, será analisada a variação na relação de concordância entre o verbo e o sujeito quando este se refere ao próprio falante, ou seja, a primeira pessoa do discurso, como exemplificado abaixo:

(1) a. Eu *trabalh**o*** na roça.

b. Eu *trabalh**a*** na roça.

(2) a. Eu *trabalh**ei*** na roça.

b. Eu *trabalh**ou*** na roça.

Para além do presente e do pretérito perfeito do indicativo — exemplificados em (1) e (2), respectivamente —, a 1ª pessoa do singular só tem um morfema próprio no futuro do presente do indicativo, tempo que praticamente não é usado na linguagem popular e rural, sendo substituído pela perífrase com o verbo *ir* funcionando como auxiliar (e.g., *vou trabalhar* por *trabalharei*). Os outros tempos verbais que têm alguma vitalidade no português popular do Brasil, como o imperfeito do indicativo e o do subjuntivo, não contêm um morfema específico para a 1ª pessoa do singular (e.g., *eu trabalhava / ele trabalhava; se eu trabalhasse / se ele trabalhasse*). Portanto, o processo de variação na concordância verbal envolvendo a 1ª pessoa do singular se circunscreve aos tempos em que essa pessoa do discurso ainda preserva uma marca própria e que têm alguma vitalidade na linguagem popular brasileira, ou seja, o presente e o pretérito perfeito do indicativo.

A hipótese de que um verbo desprovido de morfologia de pessoa e número era a forma predominante nas variedades antigas do português afro-brasileiro foi levantada por Mussa (1991, p.218-222). Embora sejam escassas as evidências empíricas diretas que possam apoiar essa proposta, existem, de fato, alguns dados

em tradições orais afro-brasileiras que apontam para a presença variável do morfema de 1ª pessoa do singular. Essa variação ocorre na lírica do *jongo*, uma forma musical afro-brasileira registrada por Ribeiro (1968, p.198-199), na década de 1950. Outros exemplos ainda figuram no *Lundu do Pai João*, uma canção afro-brasileira anônima (LOPES, 1988, p.171), e nos *vissungos*, canções afro-brasileiras registradas em São João da Chapada, Minas Gerais, na década de 1930 por Machado Filho (1985, p.80).

A ideia de que, no processo de transmissão irregular que afetou a formação do português afro-brasileiro, as marcas de pessoa e número teriam sofrido, em todas as pessoas do discurso, uma forte concorrência da forma não marcada da 3ª pessoa do singular recebe evidências favoráveis de investigações atuais sobre os processos de aquisição de segunda língua e de língua materna. Em estudos de aquisição do espanhol como segunda língua, observa-se que, na interlíngua, a 3ª pessoa do singular serve como *default* ou forma subespecificada (ANDERSON, 2002, p.91; CLEMENTS, 2003; BRUHN DE GARAVITO, 2003; McCARTHY, 2006). Também estudos da aquisição monolíngue de catalão (DAVIDSON; GOLDRICK, 2003), espanhol (RADFORD; PLOENNING-PACHECO, 1995) e português (SIMÕES; STOEL-GAMMON, 1979) apontam para o uso da 3ª pessoa do singular como *default*. Em variedades do português da África, também há evidências da variável generalização da 3ª pessoa do singular do verbo (BAXTER, 2002, p.18-19; GONÇALVES; STROUD, 1998, p.123; MENDES, 1985, p.149-151). Portanto, na população afro-brasileira em épocas passadas, seria de esperar que houvesse uma variação com a forma não marcada da 3ª pessoa do singular, podendo ser usada com referência a qualquer pessoa do discurso.

Hoje em dia, o fenômeno é raro no Brasil, mesmo considerando o português popular rural, onde a variação na flexão de número e pessoa do verbo é ampla e profunda (cf. Quadro 3). Descrições horizontais de dialetos rurais não fazem referência a esse tipo de variação. Veado (1982), que descreve a norma linguística de falantes rurais de pouca ou nenhuma escolaridade da microrregião sanfranciscana de Januária, no norte de Minas Gerais, afirma: "temos no dialeto rural apenas a flexão de 1ª pessoa do singular e de 3ª pessoa do singular, inexistindo, no geral, portanto, formas do plural" (p.56-57). Ada Rodrigues (1974, p.68-69), em sua descrição do "dialeto caipira" da região de Piracicaba, em São Paulo, afirma que "o sistema verbal apresenta duas formas no presente: uma para a 1ª pessoa do singular (*vou*) e outra para as demais (*vai*). No perfeito, há três formas: 1ª pessoa do singular (*fiz*), 3ª pessoa (*fez*), 1ª pessoa do plural (*fizemos*)". Por outra parte, a inexistência de estudos sociolinguísticos sobre a variação na concordância verbal envolvendo a 1ª pessoa do singular em normas populares urbanas aponta para a inexistência do fenômeno, mesmo considerando-se a periferia das grandes cidades. E até nas comunidades rurais afro-brasileiras isoladas estudadas no Projeto *Vertentes*

(cf. capítulo 5 deste livro), somente a comunidade de Helvécia, no Extremo Sul da Bahia, apresentou um nível significativo de variação na concordância verbal envolvendo a 1ª pessoa do singular. Assim, parece que, por um lado, a 1ª pessoa do singular foi aquela que mais resistiu à erosão provocada pelo contato do português com as línguas africanas e autóctones no Brasil; e, por outro lado, pode-se pensar que a recomposição da morfologia verbal de pessoa e número nas variedades do português afro-brasileiro teria se iniciado também pelo morfema de 1ª pessoa do singular.[3]

A existência desse processo estruturado de variação em Helvécia destaca essa comunidade no panorama linguístico do português rural do Brasil e aproxima esse dialeto dos crioulos de base lexical portuguesa da África, que eliminaram totalmente a flexão de número e pessoa existente na língua-alvo, o português. No caso de Helvécia, é mais plausível pensar que houve um violento processo de variação no uso de todos os morfemas número-pessoais, que se estendeu até as primeiras décadas do século XX. A partir daí, o uso desses morfemas se teria intensificado por influências externas à comunidade, sendo o morfema de 1ª pessoa do singular aquele que mais rapidamente se teria fixado, tanto que a variação no seu uso entre os jovens já é muito reduzida. Desse modo, o processo de implementação do uso da regra de concordância verbal que vem se desenvolvendo em Helvécia pode ser definido como um processo de mudança de tipo *descrioulizante*, pois se trata de um processo de reintrodução da morfologia da língua-alvo.

14.1.1 A variável dependente

A presente análise tem como base empírica uma amostra de fala vernácula com 18 falantes da comunidade de Helvécia agrupados em três faixas etárias: faixa 1, de 20 a 40 anos; faixa 2, de 41 a 60 anos; faixa 3, de mais de 60 anos. Os falantes estão distribuídos equitativamente pelos dois sexos, considerando-se também o seu nível de escolaridade e o fato de terem vivido, ou não, por pelo menos seis meses fora da comunidade. A análise dessas variáveis sociais fornecerá a comprovação empírica para a tendência de mudança descrioulizante aqui proposta. Para além disso, apresenta-se também um tratamento do problema do encaixamento

---

[3] Fatos semelhantes são observados na literatura sobre a aquisição do português brasileiro como língua materna. Assim, Simões e Stoel-Gammon (1979) observaram que o morfema de 1ª pessoa do singular foi o primeiro morfema de pessoa-número adquirido categoricamente, depois de uma fase em que o verbo finito apresentava a forma não marcada da 3ª pessoa do singular.

linguístico do processo de mudança, incorporando algumas variáveis estruturais. O modelo teórico-metodológico que orienta a análise é definido no quadro da Sociolinguística Quantitativa (WEINREICH; LABOV; HERZOG, 2006 [1968]; LABOV, 1972a, 1982, 1994), incorporando-se alguns elementos do Programa Gerativista (CHOMSKY, 1981, 1986, 2001; KOOPMAN; SPORTICHE, 1991) na análise das variáveis estruturais. No tratamento quantitativo dos dados, utilizou-se o pacote de programas VARBRUL (PINTZUK, 1988; SCHERRE; NARO, 2003; TAGLIAMONTE, 2006).

Dentro desse aparato metodológico, a análise estrutura-se da seguinte maneira: observa-se o comportamento da variável em questão, dita variável dependente, em função de um determinado número de variáveis estruturais e sociais, ditas variáveis independentes ou explanatórias. Os resultados gerais da variável dependente podem ser vistos na Tabela 1.

Tabela 1 - Frequência geral de aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA

VARIANTE	Nº DE OC. / TOTAL	FREQUÊNCIA
Com concordância (ex.: *eu falo / eu falei*)	2.556/3.134	82%
Sem concordância(ex.: *eu fala / eu falou)*	578/3.134	18%

O emprego do morfema de 1ª pessoa do singular com uma frequência de aproximadamente 80% do total das ocorrências analisadas, comparado com o reduzido uso do morfema de número junto à 1ª e à 3ª pessoas do plural — abaixo dos vinte pontos percentuais (cf. seções 14.2 e 14.3, adiante) —, revela que a implementação da regra de concordância verbal com a 1ª pessoa do singular está bem avançada na gramática da comunidade de fala.

14.1.2 O encaixamento linguístico

Para analisar como esse processo de variação e mudança se encaixa na estrutura linguística da comunidade, foram consideradas inicialmente as seguintes variáveis independentes ou explanatórias: (i) tipo de oração; (ii) tipo de frase; (iii) tempo do verbo; (iv) estrutura mórfica da forma verbal; (v) ordem dos constituintes na sentença; (vi) efeito de gatilho; e (vii) saliência fônica. Das variáveis estruturais propostas, o programa de regras variáveis VARBRUL selecionou, como estatisticamente significativas, pela seguinte ordem, as variáveis: (i) ordem dos constituintes na sentença; (ii) tempo do verbo; e (iii) efeito de gatilho.

A variável **ordem dos constituintes na sentença** foi organizada da seguinte maneira:

(3) ORDEM DOS CONSTITUINTES NA ORAÇÃO

a. S-V-(X) [4]

Ex.: Ieu tenho uma poção de amigo.

b. V-(X)

Ex.: Levei no médico.

c. X-S-V-(X)

Ex.: Mas, de vez em quando, eu falo com ela.

d. X-V-(X)

Ex.: Só vejo falá nome.

e. S-X-V-(X)

Ex.: Eu quase num faz não.

f. (X)-V-(X)-S-(X)

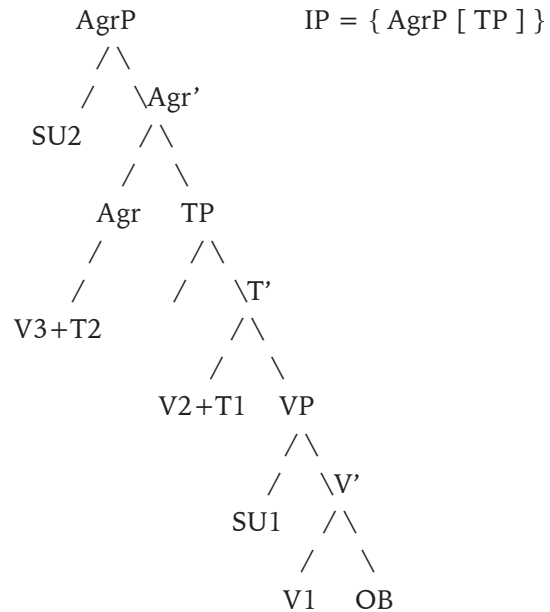
Ex.: De manhã, foi eu no jiquiá pa dá o minino cumida.

Os resultados quantitativos dessa variável são apresentados na Tabela 2.

Tabela 2 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável ordem dos constituintes na oração

ORDEM DOS CONSTITUINTES	Nº DE OC. / TOTAL	FREQ.	P.R.
S-V-(X)	893/1024	87%	**.60**
X-S-V-(X)	339/385	87%	**.59**
X-S-X-V-(X)	51/66	77%	**.48**
X-V-(X)	403/498	81%	**.47**
S-X-V (X)	57/114	50%	**.42**
V-(X)	594/793	75%	**.40**
(X)-V-(X)-S-(X)	6/11	35%	**.04**
Total	**2.556/3.134**	**82%**	—

Os seis valores dessa variável podem ser reagrupados em quatro: (i) sujeito imediatamente antes do verbo, [S-V-(X)] e [X-S-V-(X)]; (ii) sujeito nulo, [V-(X)] e [X-V-(X)]; (iii) sujeito separado do verbo por um constituinte, [S-X-V-(X)] e [X-S-X-V-(X)]; e (iv) sujeito posposto ao verbo [(X)-V-(X)-S-(X)]. A análise dessa variável fundamentou-se no diagrama estrutural apresentado no Quadro 4.

[4] S (= sujeito); V (= verbo); e X (= qualquer outro constituinte). (X) significa que este constituinte pode não ocorrer.

Quadro 4 - Representação estrutural do sintagma flexional (IP)
Fonte: Adaptado de Koopman e Sportiche, 1991.

Assim, a situação ótima para desencadear o processo de concordância seria a de sujeito lexicalmente preenchido que se move para a posição de especificador de AgrP (SU^2). O movimento do sujeito para o especificador de AgrP, associado ao movimento do verbo para tempo (V2+T1), e do núcleo tempo, que já contém o verbo, para a posição de núcleo de AgrP (V3+T2) definiriam o processo de concordância. Os pesos relativos dos fatores [S-V-(X)] e [X-S-V-(X)], quase idênticos e os mais altos do grupo, confirmam a hipótese.

Os valores relativamente mais baixos do sujeito nulo podem ser explicados na medida em que a posição de especificador de IP é preenchida por uma categoria vazia *pro*, o que, associado à fragilidade dos traços de concordância na gramática do dialeto, não favoreceria o alçamento do verbo para a posição de núcleo de AgrP, não ocorrendo, portanto, a concordância em forma fonética.

No caso de haver um constituinte entre o sujeito e o verbo, há duas possibilidades de explicação para o desfavorecimento da aplicação da regra de concordância. De um lado, pode-se pensar que, em [S-X-V-(X)], o pronome *eu* está na posição de tópico, portanto em adjunção a IP, e não na posição de especificador de AgrP. Por outro lado, para o caso de [X-S-X-V-(X)], pode-se pensar que, dada a fragilidade dos traços de concordância no dialeto, o verbo subiu para T' mas não subiu para Agr', mesmo que a posição de especificador de AgrP esteja lexicalmente preenchida pelo sujeito. Evidentemente, a presença de advérbios pré-verbais dificulta o movimento do verbo para Agr'.

Finalmente, o valor de não aplicação da regra de concordância muito elevado no fator [(X)-V-(X)-S-(X)] corrobora a hipótese de que o movimento do sujeito para a posição de especificador de AgrP é crucial para desencadear o processo de concordância.

Os resultados da variável **tempo do verbo**, que compreende o presente, o pretérito perfeito e o futuro perifrástico, são apresentados na Tabela 3.

Tabela 3 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável tempo do verbo

TEMPO DO VERBO	Nº DE OC. / TOTAL	FREQ.	P.R.
Presente do Indicativo	1452/1739	83%	**.56**
Pretérito Perfeito do Indicativo	1009/1297	78%	**.39**
Futuro Perifrástico	95/98	97%	**.88**
Total	**2.556/3.134**	**82%**	—

Os resultados indicam que a implementação do uso da regra de concordância se dá primeiramente no presente e depois no pretérito. Por outro lado, o alto valor do futuro perifrástico (*eu vou fazer ~ eu vai fazer*) pode ser explicado pelo fato de o auxiliar ser, em realidade, uma manifestação da categoria funcional de tempo. Portanto, o auxiliar já seria gerado na posição de núcleo do TP, o que favorece bastante o seu alçamento para o núcleo de Agr, estabelecendo-se a concordância. Além disso, a alta frequência desse auxiliar e a sua saliência fônica também facilitam a implementação do morfema de pessoa.

Os resultados da variável **efeito de gatilho** indicam que os fatores de natureza interativa e discursiva também interferem no desempenho linguístico dos informantes. Nesse grupo, foram definidos dois valores: efeito desfavorável e sem efeito, exemplificados abaixo:

(4) EFEITO DE GATILHO

a. efeito desfavorável

Ex.: — E onde você *trabalha*?

— *Trabalha* cum tratores.

b. sem efeito

Ex.: — E tem quanto tempo que isso aconteceu?

— Isso, num me *engano*, acho tem, mais ou menos, uns... uns quatro ano.

Em (4a), a forma não marcada da 3ª pessoa do singular na pergunta do documentador pode ter induzido o falante a não fazer a concordância em sua resposta. Já uma pergunta como a de (4b) não interferiria no fenômeno em foco. Os resultados do processamento quantitativo dos dados dessa variável são apresentados na Tabela 4.

Tabela 4 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável efeito de gatilho

EFEITO DE GATILHO	Nº DE OC. / TOTAL	FREQ.	P.R.
Efeito desfavorável	268/383	70%	**.31**
Sem efeito	2288/2751	83%	**.53**
Total	**2.556/3.134**	**82%**	—-

No caso do efeito desfavorável, a pergunta feita pelo entrevistador contém uma forma verbal de 3ª pessoa do singular (relativa a um tratamento de *você* ou *o/a senhor/a*), e isso induz o informante a empregar, em sua resposta, o verbo na 3ª pessoa do singular, mesmo quando está se referindo a si mesmo. Isso se reflete nos resultados percentuais e dos pesos relativos, pois a frequência de aplicação da regra de concordância cai, nesse contexto, de 82% (frequência geral) para 70%, com peso relativo final de .31.

14.1.3 O encaixamento social

No âmbito das variáveis sociais, os resultados da análise quantitativa dos dados revela um quadro notavelmente límpido, de tal maneira que todas as variáveis sociais propostas foram selecionadas como significativas em termos estatísticos pelo programa de regras variáveis. Desse modo, os resultados das variáveis sociais, particularmente a idade, indicam claramente um processo de mudança em curso descrioulizante de implementação da regra de concordância verbal.

Dentro da literatura sociolinguística, a dicotomia variação estável *versus* mudança em curso define-se primariamente a partir da **variável idade** (LABOV, 1981; LUCCHESI, 2001b). No caso da variação estável, verifica-se um padrão curvilinear, no qual as faixas intermediárias exibem os valores mais elevados da variante de prestígio. Já para a mudança em curso, observa-se um padrão ascendente, no qual os mais jovens apresentam os valores mais elevados da variante inovadora. No caso dos resultados da variável faixa etária dessa amostra de fala da comunidade de Helvécia, pode-se mesmo pensar numa reta, pois há uma simetria entre a faixa 3 (mais de 60 anos), com praticamente 30 pontos abaixo da marca de neutralidade, e a faixa 1 (falantes de 20 a 40 anos) com 30 pontos acima, o que define o padrão linear ascendente, característico da mudança em curso, como se pode ver na Tabela 5 e na Figura 1.

Tabela 5 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável faixa etária

FAIXA ETÁRIA	Nº DE OC. / TOTAL	FREQ.	P.R.
20 a 40 anos	1017/1060	96%	**.80**
40 a 60 anos	787/920	86%	**.51**
Mais de 60 anos	752/1154	65%	**.21**
Total	**2.556/3.134**	**82%**	—

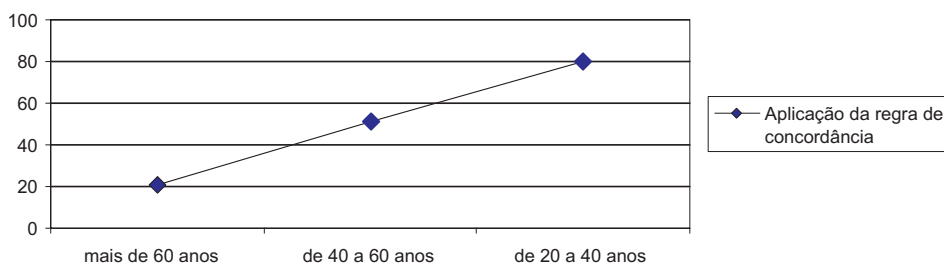



Figura 1 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável faixa etária (pesos relativos)

Ressalve-se que essa foi a primeira variável selecionada pelo programa dentre as estatisticamente significativas.

As demais variáveis sociais apontam também no sentido de uma mudança em curso em direção à implementação da regra de concordância, proveniente de influxos externos à comunidade de fala. No caso da **variável sexo**, como se pode ver na Tabela 6, os homens lideram o processo de mudança.

Tabela 6 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável sexo

SEXO	Nº DE OC. / TOTAL	FREQ.	P.R.
Homens	1.211/1.474	82%	**.56**
Mulheres	1.345/1.660	81%	**.44**
Total	**2.556/3.134**	**82%**	—

Os homens lideram o processo de mudança em função do seu maior contato com o mundo exterior. Enquanto as mulheres se circunscrevem ao universo doméstico, dividindo com o seu companheiro apenas o trabalho na roça, os homens são em geral os responsáveis pela comercialização dos pequenos excedentes da produção agrícola familiar na feira da cidade e aqueles que mais se deslocam para os trabalhos sazonais nos grandes centros urbanos. Portanto, são eles os primeiros a incorporar o uso das variantes do padrão urbano. Isso fica mais nítido no resultado dos pesos relativos, com .56 para os homens frente aos .44 para as mulheres.

A influência do exterior é mensurada de forma mais direta com a **variável estada fora da comunidade** por um período não inferior a seis meses, cujos resultados são apresentados na Tabela 7.

Tabela 7 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável estada fora da comunidade

ESTADA FORA DA COMUNIDADE	NºDE OC. / TOTAL	FREQ.	P.R.
Pelo menos seis meses	1.034/1.109	93%	**.67**
Não	1.522/2.025	75%	**.41**
Total	**2.556/3.134**	**82%**	—-

Também nesse caso, os falantes que moraram fora da comunidade lideram o processo de mudança, o que se reflete, tanto na frequência de uso (93%, para quem viveu fora, contra 75%, para quem nunca viveu fora da comunidade), quanto nos pesos relativos (.67 *versus* .41). Aqueles que já viveram em centros urbanos puderam assimilar diretamente a variante típica do padrão urbano culto.

Surpreendentes foram os resultados da **variável escolaridade**, apresentados na Tabela 8, já que os pesos relativos invertem os percentuais de frequências absolutas, de modo que os falantes analfabetos apresentam um peso relativo maior para aplicação da regra de concordância que os falantes semialfabetizados.

Tabela 8 - Aplicação da regra de concordância verbal com a 1ª pessoa do singular em Helvécia-BA segundo a variável nível de escolaridade

NÍVEL DE ESCOLARIDADE	Nº DE OC. / TOTAL	FREQ.	P.R.
Semianalfabeto	662/740	89%	**.39**
Analfabeto	1.894/2.394	79%	**.53**
Total	**2.556/3.134**	**82%**	—

Apesar da possibilidade de que um possível enviezamento dos dados tenha interferido no processamento quantitativo, deve-se destacar a precariedade do sistema de educação pública na zona rural. Tanto é assim que os falantes considerados como semianalfabetos possuíam apenas a habilidade de "desenhar" o próprio nome. Assim, a distinção foi feita apenas entre semianalfabetos, que tinham qualquer experiência de escolaridade (mesmo que disso resultasse apenas a habilidade para assinar o nome) e analfabetos plenos. Como se pode ver, tal distinção não foi de monta a interferir no comportamento linguístico dos falantes no sentido do que seria o esperado.

14.1.4 Conclusão

O encaixamento linguístico do uso do morfema verbal de 1ª pessoa do singular na gramática da comunidade rural afro-brasileira de Helvécia revelou que a concordância verbal é favorecida quando o sujeito pronominal está realizado imediatamente antes do verbo e é muito desfavorecida quando o sujeito está posposto ao verbo. Essa é uma correlação que se aplica ao fenômeno da concordância verbal como um todo no português brasileiro, independentemente da pessoa do discurso, e se observa também em outras línguas, tornando-se um sério candidato à condição de ser um condicionador universal do processo de variação e mudança em relação à regra de concordância verbal, no âmbito do que Weinreich, Labov e Herzog (2006 [1968]) definiram como *problema das restrições* (ing. *constraints problem*). No que concerne especificamente à aplicação da regra de concordância verbal com

a primeira pessoa do discurso, a aquisição do morfema flexional ocorre primeiramente no presente do indicativo, para depois ocorrer no pretérito perfeito. Já no âmbito da interação conversacional, a presença de uma forma verbal não marcada quanto à pessoa na pergunta posta pelo interlocutor tende a inibir o emprego da forma verbal flexionada na resposta do falante.

Os resultados do efeito das variáveis sociais revelaram que o dialeto de Helvécia se encontra em um estágio avançado do processo de aquisição da regra de concordância verbal com a 1ª pessoa do singular. Nesse sentido, foram particularmente significativos os resultados obtidos na variável faixa etária, pois indicaram claramente a direção da mudança no sentido da aquisição da regra, configurando, prospectivamente, o processo de implementação e, retrospectivamente, um longo processo diacrônico, através do qual um sistema anterior praticamente sem flexão de número e pessoa foi sendo substituído em um processo de nivelamento linguístico a partir de influências externas à comunidade. Outras variáveis sociais (sexo, estada fora da comunidade) confirmaram os resultados da variável faixa etária, definindo uma tendência iminente de homogeneização dos padrões linguísticos da comunidade de fala em torno do novo sistema com a concordância verbal. O reduzido nível de variação no uso do morfema flexional da 1ª pessoa do singular entre os habitantes de Helvécia na faixa dos 20 a 40 anos (menos de 5%) indica que em breve o fenômeno deixará de ter significância estrutural na gramática da comunidade de fala, tornando-se a variação residual um fenômeno de desempenho. Com isso, uma das marcas mais notáveis da influência do contato entre línguas sobre esse dialeto terá desaparecido — um processo que pode ter ocorrido em inúmeras comunidades quilombolas brasileiras, quando essas se integraram na sociedade brasileira ao longo do século XX, ou foram simplesmente desarticuladas em função da violenta penetração do capitalismo no campo brasileiro.

Portanto, os resultados da análise sociolinguística que aqui se apresentam são particularmente interessantes, porque revelam que o contato entre línguas na formação de comunidades rurais brasileiras chegou a afetar o emprego dos morfemas flexionais do verbo em todas as pessoas do discurso, em um processo que se assemelha bastante ao que ocorre nos casos típicos de crioulização, em que toda a morfologia de número e pessoa do verbo eventualmente presente na língua-alvo é suprimida. Desse modo, o processo de variação que se observa em Helvécia está muito mais próximo daqueles que afetaram a língua portuguesa no arquipélago de Cabo Verde[5] e nas ilhas de São Tomé e Príncipe, do que qualquer suposto processo

[5] É de notar-se que há dialetos do crioulo de Cabo Verde que, além de apresentar um verbo predominantemente desprovido de morfologia flexional, mantêm em alguns verbos a flexão de tempo e aspecto, mas não a flexão de pessoa e número. Já o crioulo de Damão (na Índia) é o que mais se assemelha formalmente aos fatos comentados aqui, pois o verbo nesse crioulo não tem flexão de pessoa e número, mas tem flexão de tempo e aspecto (BAXTER, 1997).

de variação que tenha acometido a língua portuguesa em Portugal desde a formação do reino português até os dias atuais.

Por outro lado, pode-se pensar em um quadro de vários estratos diferenciando as variedades linguísticas brasileiras, mesmo aquelas do interior do país. Assim, só nas comunidades de fala mais profundamente afetadas pelo contato entre línguas, o processo de erosão gramatical atingiu todas as pessoas do discurso. Nas demais variedades populares do português brasileiro, que ainda assim foram significativamente afetadas pelo contato entre línguas em sua formação histórica, todos os morfemas número-pessoais foram atingidos, exceto o morfema que indica a pessoa do próprio falante, o morfema da auto-referência. Na norma urbana culta, em contraste, observa-se apenas um pequeno processo de variação que afeta somente o emprego do morfema de 3ª pessoa do plural, que no mais das vezes se atualiza pelo débil processo fônico de nasalização de uma vogal átona final. Tal contraste ratifica, por um lado, a polarização sociolinguística do Brasil, nos termos propostos por Lucchesi (1994, 1998, 2001a, 2002a, 2002b, 2006). De outra perspectiva, reúnem-se elementos suficientes para estabelecer a seguinte correlação empiricamente motivada no cenário sociolinguístico e histórico do português brasileiro: quanto mais ligada a situações de contato entre línguas em sua história estiver uma comunidade de fala brasileira, maior será hoje o grau de variação no uso de seus morfemas flexionais.

Por fim, a grande resistência do morfema de 1ª pessoa do singular nesse amplo processo de variação e mudança que se observa no PB forneceria elementos para a formulação de mais uma questão no plano do que Weinreich, Labov e Herzog (2006 [1968]) definiram como *problema das restrições*, ou seja, a problemática de se definirem condicionamentos (ou restrições) universais da variação e da mudança linguísticas. O cenário brasileiro induziria o analista a postular que o morfema referente à pessoa do próprio falante seria o mais refratário aos processos de variação e mudança que afetassem o paradigma de flexão verbal de pessoa e número. Entretanto, uma rápida consideração de outras línguas da família indo-europeia que também passaram por processos de variação e mudança em relação ao seu sistema de concordância verbal, como o inglês e o francês, oferece de pronto contraexemplos que negam tal hipótese. No inglês, o único morfema remanescente é o da 3ª pessoa do singular (curiosamente o morfema da não pessoa). No francês, os morfemas de maior vitalidade atualmente são os da 1ª e da 2ª pessoas do plural. Portanto, ao que tudo indica, os processos de variação e mudança que afetam a flexão verbal de número e pessoa nas línguas humanas não exibem qualquer escala implicacional ou hierárquica no que concerne às pessoas do discurso.

De certo mesmo, apenas a correlação empiricamente motivada entre os processos de transmissão linguística irregular e a erosão dessa morfologia flexional de pessoa e número, como se pode comprovar com análises objetivas da ampla

variação que se observa no PB, no contexto de um processo de transmissão linguística irregular mais leve, em cotejo com a eliminação de toda a flexão de pessoa e número do verbo, que se observa em todos os crioulos de base portuguesa da África e da Ásia, que passaram por um processo mais intenso e radical de transmissão linguística irregular.

14.2 A concordância verbal com a terceira pessoa do plural

No que diz respeito à variação na aplicação da regra de concordância verbal associada à 3ª pessoa do plural, o panorama sociolinguístico da língua portuguesa no Brasil fornece importantes evidências empíricas favoráveis à hipótese do contato entre línguas como móvel do processo de variação e mudança.[6] A primeira delas diz respeito aos níveis diferenciados de variação que se encontram nas diferentes variedades do português brasileiro, constituindo o seguinte *continuum*:

(i) nas comunidades rurais afro-brasileiras do interior do Estado da Bahia, o nível de aplicação da regra de concordância é da ordem de 16% (ver abaixo);

(ii) em comunidades de pescadores analfabetos ou pouco escolarizados, no norte do Estado do Rio de Janeiro, esse nível sobe para 38% (VIEIRA, 1997);

(iii) na fala de analfabetos da cidade do Rio de Janeiro, o nível de aplicação já é de 48% (NARO, 1981);

(iv) na fala de indivíduos escolarizados das cidades do Rio de Janeiro e de Florianópolis, na região Sul do Brasil, a regra de concordância verbal já é aplicada com uma frequência de 73% e 79%, respectivamente (SCHERRE; NARO, 1997; MONGUILHOTT; COELHO, 2002, respectivamente);

(v) na norma urbana culta do Rio de Janeiro, os falantes usam a regra com uma frequência de 94% (GRACIOSA, 1991).

O cotejo dos dois extremos do *continuum* constitui prova contundente da polarização sociolinguística do Brasil: enquanto o uso da regra é quase categórico entre falantes urbanos com elevado grau de escolaridade, o nível de variação é elevadíssimo nas comunidades rurais afro-brasileiras, como demonstrarão os resultados da análise variacionista que serão apresentados nesta seção. Por outro lado, confirma-se a relação entre o contato entre línguas e o grau de variação linguística, na medida em que as comunidades rurais exibem um grau maior de

[6] Esta seção é baseada nos resultados da pesquisa de Mestrado de Jorge Augusto Alves da Silva (2003), orientada por Dante Lucchesi.

variação do que as comunidades urbanas, já que o contato entre línguas teria predominado no interior do país (cf. capítulo 1 deste livro); e, entre as comunidades rurais, o nível de variação é maior nas comunidades afro-brasileiras, que têm a sua origem mais diretamente ligada ao contato do português com as línguas africanas (cf. capítulo 2 deste livro).

Outra evidência empírica relevante são as tendências divergentes dos processos de variação e mudança observados nas diferentes comunidades. Enquanto nas comunidades urbanas os resultados apontam para um cenário de variação estável (SCHERRE; NARO, 1997, p.107), ou um ligeiro declínio na aplicação da regra entre os mais jovens (GRACIOSA, 1991, p.71), nas comunidades rurais a tendência é de incremento do uso da regra de concordância, refletida na curva ascendente dos resultados da variável faixa etária, em que os jovens são aqueles que mais aplicam as regras de concordância, enquanto os mais velhos são os que menos fazem a concordância verbal (VIEIRA, 1997, p.128-130; e a análise abaixo). Assim, confirma-se a hipótese de que, na situação de contato em que esses dialetos se formaram, as regras de concordância teriam sido profundamente afetadas; ao passo que, a partir de meados do século XX, a regra de concordância estaria sendo reintroduzida nessas comunidades rurais por influência dos modelos linguísticos urbanos, através do deslocamento populacional, da influência dos meios de comunicação de massa e da massificação do ensino público. Esses resultados também negam a hipótese de uma deriva secular, proposta por Naro e Scherre (1993, 2007), que determinaria uma contínua e progressiva eliminação da morfologia flexional do verbo.

A análise variacionista da aplicação da regra de concordância verbal junto à 3ª pessoa do plural que aqui se apresenta teve como universo de observação três comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia: Cinzento, no Município de Planalto; Helvécia, no Município de Nova Viçosa; e as comunidades geminadas de Barra e Bananal, no Município de Rio de Contas. Foram analisadas 12 entrevistas, com cerca de 50 minutos de duração cada uma, para cada comunidade. Essas entrevistas foram realizadas segundo as técnicas da pesquisa sociolinguística com falantes com pouca ou nenhuma escolarização, distribuídos equitativamente entre os dois sexos e em três faixas etárias (cf. capítulo 5 deste livro).

14.2.1 O encaixamento linguístico

Na amostra de fala assim constituída, foram depreendidas 1.706 formas verbais ligadas a sujeitos da 3ª pessoa do plural. A regra de concordância foi aplicada em 273 ocorrências, correspondendo a 16% do total. Os exemplos (5) e (6) abaixo

ilustram a variação na aplicação da regra de concordância verbal com a 3ª pessoa do plural.[7]

(5) Eles **ganha** pouco. (ptg. padrão: Eles ganham pouco.)

(6) As mulhé **vai**. (ptg. padrão: As mulheres vão.)

A comunidade de Cinzento apresentou o menor percentual de concordância (13%), enquanto Rio de Contas demonstrou o maior percentual (24%). Em Helvécia, o percentual de concordância ficou em 16%. Esses resultados fogem um pouco à expectativa criada a partir das informações disponíveis sobre a formação de cada uma dessas comunidades (cf. capítulo 2 deste livro): a comunidade de Helvécia, para a qual se tem um registro direto de um passado anterior crioulizante, apresentaria a menor frequência de aplicação da regra de concordância. Isso mostra a complexidade de uma relação extremamente mediada entre os padrões coletivos de comportamento linguístico de uma comunidade de fala e os processos sócio-históricos em que esses padrões se formaram.

A análise variacionista revelou que seis fatores estruturais condicionam a aplicação da regra de concordância na 3ª pessoa do plural nas comunidades de fala analisadas: (i) saliência fônica; (ii) forma de indicação do plural no sujeito; (iii) concordância nominal no SN sujeito; (iv) realização e posição do sujeito; (v) caracterização semântica do sujeito; e (vi) tipo de verbo.

O **princípio da saliência fônica**, proposto para o português por Naro e Lemle (1976) e que tem sido aplicado por praticamente todos os estudos variacionistas que trataram desse tema (NARO, 1981; VIEIRA, 1997; MONGUILHOTT, 2001; entre outros), serviu para explicar o curso da aquisição da regra de concordância, pois a hierarquia de saliência foi respeitada, demonstrando que, com formas verbais que apresentam maior diferenciação morfofonológica na oposição singular-plural, a frequência de aplicação da regra é maior do que com formas verbais que apresentam pouca diferenciação morfofonológica para marcar o plural, como se pode ver na Tabela 9 e na Figura 2.

Tabela 9 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável saliência fônica

SALIÊNCIA FÔNICA	Nº DE OC. / TOTAL	FREQ.	P.R.
1. nível baixo de saliência fônica (ex.: sai/saem; bate/batem; fala/falam)	48/818	6%	**.27**
2. nível intermediário de saliência (ex.: faz/fazem; tá/tão; bateu/bateram; quer/querem; vai/vão; foi/foram)	158/675	23%	**.69**
3. flexão com nível alto de saliência (ex.: quis/quiseram; fez/fizeram; é/são; veio/vieram)	67/213	31%	**.78**

[7] Todos os exemplos apresentados nesta seção são retirados da base de dados depreendida da amostra de fala analisada. Graficamente, o sujeito será marcado pelo sublinhado e o verbo em foco estará em negrito.

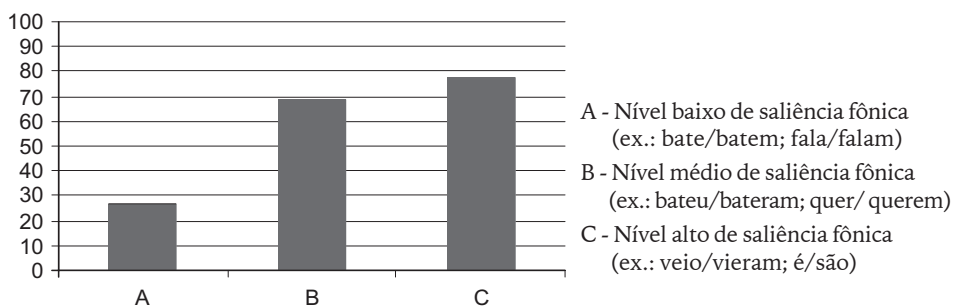



Figura 2 - Nível de concordância, segundo o grau de saliência fônica do verbo

De acordo com a visão de Guy (2005, p.26), "na aquisição de uma regra por uma pessoa que não tem essa regra, a escala da saliência atuaria no sentido de que ela adquiriria primeiro os casos mais óbvios, mais salientes, e os mais difíceis de aprender seriam adquiridos só depois". Dessa forma, os resultados da variável saliência fônica podem ser interpretados como uma evidência de que está em curso, nessas comunidades rurais afro-brasileiras isoladas, um processo de mudança aquisicional da regra de concordância verbal.

A segunda variável selecionada pelo programa foi a **forma de indicação do plural no SN sujeito**. Esta variável foi estruturada com três valores: (i) indicação mórfica do plural, com o emprego do morfema de número *-s*; (ii) indicação do plural com o reforço de um numeral ou quantificador; (iii) indicação do plural através de um substantivo coletivo, o *plural lexical;*[8] esses fatores são exemplificados, respectivamente, a seguir:

(7) Os cara **tiraro** trinta pessoa.

(8) Mutchos **chamavam** até de urubu.

(9) A nova geração **tão** animado.

A Tabela10 apresenta os resultados do processamento quantitativo dos dados.

Tabela 10 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável indicação do plural no SN sujeito

INDICAÇÃO DO PLURAL NO SN SUJEITO	Nº DE OC. / TOTAL	FREQ.	P.R.
Mórfica	135/1131	12%	**.47**
Numeral ou quantificador	18/103	17%	**.52**
Lexical	24/30	80%	**.97**

Os resultados revelam que a simples marca mórfica no núcleo do sujeito ou no seu determinante desfavorece a aplicação da regra de concordância verbal. O fator a mostrar maior influência na pluralização do verbo foi o lexical, isto é, um

[8] O coletivo, então, não possuiria marca de plural, mas a ideia de plural.

vocábulo sem marca formal de plural, mas com a ideia de plural levou o verbo, na maioria dos casos em que apareceu, a flexionar-se. O quantificador e o numeral favoreceram mais a aplicação da regra do que a marcação mórfica. A pluralização com o sujeito coletivo é mais do que um fato morfossintático, já que é um fenômeno semântico, pois o que vem primeiro à mente do falante é a ideia de plural, de mais de um indivíduo, portanto, a indicação do plural, nesses casos, é mais transparente. Dessa forma, agiria a força da *transparência semântica* a determinar a pluralização.[9]

Uma outra variável estrutural que se revelou interessante foi a correlação entre a **concordância nominal no SN sujeito** e a concordância verbal. Observou-se que a probabilidade de o falante aplicar a regra de concordância verbal era muito maior quando o falante aplicava a regra de concordância nominal de número no SN sujeito do verbo, como exemplificado na frase abaixo:

(10) A**s** pesso**as** custuma**m** comprar muito.

A Tabela 11 contém os resultados do processamento quantitativo dos dados.

Tabela 11 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável concordância nominal no SN sujeito

CONCORDÂNCIA NOMINAL NO SN SUJEITO	Nº DE OC. / TOTAL	FREQ.	P.R.
Com concordância de número	10/41	24%	**.74**
Sem concordância de número	55/601	9%	**.48**

Os resultados confirmam a hipótese inicial, já que, quando o falante deixa de fazer a concordância de número no SN sujeito, a frequência de aplicação da regra de concordância verbal cai de 24% para 9%, o que se reflete nos valores dos pesos relativos: .74 *versus* .48. Isso pode ser explicado pelo *princípio da coesão estrutural* proposto por Lucchesi (cf. seção 4.2.2. do capítulo 4 deste livro). Sumariamente, esse princípio prevê que, numa situação em que há concorrência entre duas gramáticas na competência linguística do falante, o predomínio de uma gramática sobre a outra se dá em porções determinadas da estrutura da frase. No caso das comunidades rurais afro-brasileiras isoladas, observa-se a concorrência entre uma gramática mais crioulizante (sem regras de concordância), ligada às origens históricas dessas comunidades, e uma gramática mais padrão (com regras de concordância), que estaria penetrando na comunidade em função das influências externas. Essa concorrência se manifesta na produção verbal de cada falante da comunidade, sendo restringida pelo princípio da coesão estrutural, de modo que, por exemplo, se o falante inicia uma oração com a gramática padrão (o que implica

[9] O *princípio da transparência semântica* desenvolvido na crioulística indica que as regras gramaticais se conservam mais ou são mais rapidamente adquiridas nas estruturas em que a relação entre a forma e o significado é mais direta, isto é, mais *transparente* (SEUREN; WEKKER, 1986).

a aplicação da regra de concordância nominal no SN sujeito), ele tende a aplicar a regra de concordância também na relação entre esse sujeito e o verbo, pois é o sujeito que especifica a concordância verbal. Assim, uma frase do tipo *a**s** pessoa**s** compraØ* seria mais improvável.

O princípio da coesão estrutural também se refletiu nos resultados da variável **realização e posição do sujeito**. Essa variável foi estruturada com quatro fatores: (i) sujeito anteposto, (ii) sujeito posposto, (iii) sujeito não realizado e (iv) sujeito retomado por um pronome relativo, exemplificados, respectivamente, abaixo:

(11) Eles dois **chegaro** no fim.

(12) **Nasce** otros tipo de capim.

(13) É. Eles num tão preocupado com isso não, moço, viu? Aí... num **tão** ligano pra isso não, moço.

(14) Tem uns que **pede** músiga.

A Tabela 12 mostra que o maior índice de realização da concordância verbal se dá exatamente quando o sujeito não está realizado.

Tabela 12 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável realização e posição do sujeito

REALIZAÇÃO E POSIÇÃO DO SUJEITO	Nº DE OC. / TOTAL	FREQ.	P.R.
Sujeito não realizado	80/297	27%	**.61**
Sujeito posposto	10/94	11%	**.51**
Sujeito anteposto	158/1134	14%	**.49**
Pronome relativo	13/148	9%	**.35**

Nos crioulos típicos, a ausência de flexão de número e pessoa nos verbos torna obrigatória a presença do pronome sujeito com valor referencial, de modo que a oração destacada no exemplo (13) acima teria obrigatoriamente um pronome sujeito (*eles*) realizado foneticamente, se fosse dita numa língua crioula. Pode-se mesmo pensar que o aumento do uso do pronome sujeito no português brasileiro, demonstrado por Duarte (1995), está associado à perda da flexão de número e pessoa do verbo,[10] que, por sua vez, tem em suas origens os efeitos do contato entre línguas. Portanto, se o falante omite o pronome sujeito, pode-se inferir que a gramática padrão (não crioulizante) está predominando na produção dessa oração, de modo que aumenta a probabilidade de uso da regra de concordância verbal de acordo com o princípio da coesão estrutural.

Por outro lado, não se observou uma diferença significativa no que concerne à posição do sujeito realizado. O sujeito posposto até superou ligeiramente o

[10] Quint (2008, p.77-78) apresenta uma correlação entre a perda da flexão verbal de número e pessoa e a realização obrigatória do pronome sujeito nas línguas galo-românicas.

sujeito anteposto em termos de pesos relativos (.51 contra .49). Esse resultado contraria a grande maioria dos resultados apresentados para essa variável, como se pode ver nas análises de Vieira (1997, p.126-127), Naro e Scherre (1997, p.102-103) e Monguilhott (2001, p.198-199). A explicação para esse fato pode estar no caráter descrioulizante da mudança verificada nas três comunidades. A posposição do sujeito não é um traço comum das línguas crioulas típicas, cuja ordem tende a ser do tipo SVO rígida. Haveria, portanto, a influência da língua-alvo atuando, não só na posposição do sujeito, como também na tendência de levar o verbo ao plural nessa posição, considerando-se mais uma vez o princípio da coesão estrutural.

Por fim, observou-se também que sujeito retomado por um pronome relativo é o contexto que mais inibe a aplicação da regra de concordância. Esse resultado ajusta-se à proposta de Ilza Ribeiro, apresentada no capítulo 7 deste livro, segundo a qual o relativizador em uma oração relativa com antecedente não teria uma natureza pronominal, e sim a de um complementizador neutro, desprovido dos traços de pessoa e número. Por não possuir tais traços, o relativizador não seria, em princípio, capaz de desencadear o processo de concordância verbal.

Já os resultados da variável **caracterização semântica do sujeito** conformaram-se aos resultados que predominam nas análises variacionistas sobre o tema. Tradicionalmente, o traço semântico [+humano] — cf. exemplo (15) abaixo — tende a favorecer a aplicação da regra de concordância, enquanto o traço semântico [–humano] — cf. exemplo (16) abaixo — tende a desfavorecer a sua aplicação.

(15) Algumas pessoa **dero**.

(16) Os carro **chegô** tudo cheio.

Os resultados da análise variacionista da fala das comunidades rurais afro-brasileiras confirmaram esse princípio, como se pode ver na Tabela 13.

Tabela 13 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável caracterização semântica do sujeito

TRAÇO SEMÂNTICO DO SUJEITO	Nº DE OC. / TOTAL	FREQ.	P.R.
[+humano]	245/1428	17%	**.52**
[–humano]	28/250	10%	**.40**

O favorecimento da concordância verbal quando o sujeito se refere a seres humanos está ligado à ideia de *agente*. Aquele que pratica a ação relaciona-se diretamente à ação, logo o verbo tenderia a concordar mais com ele.

A última variável estrutural a ser considerada nesta análise é o **tipo de verbo**, que foi estruturada a partir de quatro fatores: (i) verbos intransitivos, (ii) verbos transitivos, (iii) verbos auxiliares e (iv) verbos inacusativos, ilustrados, respectivamente, nos exemplos abaixo:

(17) As mulhé não **anda**, né?

(18) Eles **via** os negro, procurava ficá mais separado, né?

(19) Eles **tão** passeando.

(20) Só **fica** as gáia intêra.

Os resultados dessa variável, apresentados na Tabela 14, também estariam relacionados com o caráter agentivo do sujeito.

Tabela 14 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável tipo de verbo

TIPO DE VERBO	Nº DE OC. / TOTAL	FREQ.	P.R.
Intransitivo	64/396	16%	**.58**
Transitivo	104/688	15%	**.51**
Auxiliar	97/527	18%	**.47**
Inacusativo	8/95	8%	**.29**

Os fatores que mais favoreceriam a concordância verbal seriam os verbos intransitivos e transitivos; um pouco abaixo desses últimos ficariam os verbos auxiliares; enquanto os verbos inacusativos constituiriam o fator que mais inibiria a aplicação da regra de concordância. Note-se que os verbos intransitivos e transitivos tendem a selecionar, no léxico, em sua grade argumental um *agente*, que se realiza na sintaxe como sujeito na estrutura argumental do tema (CHOMSKY, 1986). O caráter agentivo do sujeito junto a esses verbos poderia explicar a sua atuação como um fator que favorece a concordância verbal. Já com os verbos inacusativos ocorre o contrário, sua grade temática prevê apenas um *tema/paciente* que, em vez de se realizar como objeto direto (OD) do verbo na sintaxe, ocupa a posição de sujeito em sua estrutura argumental (donde a designação de inacusativos a esses verbos, em função da sua incapacidade de atribuir caso acusativo ao constituinte que seria o seu OD, de modo que esse constituinte ocorreria como sujeito, recebendo caso nominativo, através da concordância verbal).[11] Esse caráter não agentivo dos sujeitos dos verbos ditos inacusativos explicaria a baixa frequência da concordância verbal nesse contexto, bem como a maior probabilidade de inversão verbo-sujeito que também se verifica com esses verbos.

14.2.2 O encaixamento social

A análise do encaixamento social da variação na concordância verbal junto à 3ª pessoa do plural nas comunidades de fala analisadas foi estruturada, inicialmente, com base em cinco variáveis sociais: sexo, faixa etária, estada fora da

[11] Cf. CHOMSKY, 1986.

comunidade (por pelo menos seis meses), nível de escolaridade e comunidade. Essa última variável assenta na necessidade de apontar, dentro das três comunidades estudadas, a relação das características sócio-históricas de sua formação com o processo variável em foco.[12] Das variáveis propostas, o programa de regras variáveis VARBRUL selecionou como estatisticamente relevantes as seguintes: faixa etária, comunidade e sexo.

Os resultados da análise quantitativa da variável **faixa etária** apontaram para um processo de mudança em curso, com o incremento da aplicação da regra de concordância na fala dos mais novos, numa proporção significativa, como se pode constatar nos valores da Tabela 15.

Tabela 15 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável faixa etária

FAIXA ETÁRIA	Nº DE OC. / TOTAL	FREQ.	P.R.
Faixa 1 (20 a 40 anos)	141/634	22%	**.62**
Faixa 2 (41 a 60 anos)	85/602	14%	**.48**
Faixa 3 (61 em diante)	47/470	10%	**.36**

A tendência de aplicação da regra verificada nos falantes da faixa 1 demonstra que a mudança em curso é um processo decorrente da ampliação de horizontes desse grupo, pois ele está mais em contato com os grupos externos e sente a pressão que deles procede. Além disso, eles são os mais atingidos pelos meios de comunicação, pela televisão e pelo rádio, que imprimem valores diferentes dos estabelecidos pela comunidade. Os mais velhos, por outro lado, tendem a apresentar um caráter conservador, especialmente as mulheres. Elas se inclinam a reproduzir os padrões da língua aprendida no círculo doméstico, espaço legítimo para a mulher na zona rural. O pouco contato com outros grupos, certamente, não permite que elas sejam expostas à outra realidade que poderia alterar-lhe a forma de ver o mundo e de interagir com ele.

Pode-se pensar, assim, que os mais velhos apresentam um comportamento linguístico e social mais conservador, mantendo as tradições da comunidade. No caso específico de Cinzento, os seus membros prezam os valores autóctones como forma de manter a identidade do grupo, sendo essa uma preocupação consciente ou inconsciente dos mais velhos. Uma das manifestações de Cinzento, os benditos de Santo José, reflete essa visão de conservação, pois sempre é o membro mais velho, no caso Ana Isidora (107 anos), o responsável por "puxar" o bendito e levá-lo até o fim; na ausência do mais velho, a ordem etária sucessiva obriga o próximo a entoar primeiro o canto e ser seguido pelos demais.

[12] Uma caracterização da formação sócio-histórica de cada uma das comunidades pode ser encontrada no capítulo 2 deste livro.

É interessante observar também que o estudo de Vieira (1997) aponta para o mesmo processo de aumento de uso da regra de concordância em comunidades rurais não marcadas etnicamente. Nesse sentido, pode-se pensar que, nas normas vernáculas rurais, observa-se uma mudança de cima para baixo, com o incremento do uso das regras de concordância.

A história da constituição de cada comunidade muito tem a dizer sobre as especificidades de tal processo, como ficou demonstrado com os resultados da variável **comunidade**. Como se pode ver nos resultados da Tabela 16, o curso de aquisição da concordância encontra-se em estado mais avançado na comunidade de Rio de Contas, pois o grande fluxo turístico que vivencia o município e as regiões circunvizinhas tem atraído um enorme contingente de pessoas que trazem consigo os padrões culturais da vida urbana. Já a comunidade de Cinzento foi a que exibiu os menores índices de concordância, em função de sua condição precária com poucos atrativos para os forasteiros, o que a mantém em uma situação de maior isolamento. Helvécia, por sua vez, tem vivenciado uma transformação motivada pelo fluxo de pessoas que saem do povoado e retornam e pela ação do plantio de eucalipto, o que a coloca num nível intermediário quanto à aplicação da regra de concordância verbal.

Tabela 16 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável comunidade

COMUNIDADE	Nº DE OC. / TOTAL	FREQ.	P.R.
Rio de Contas	98/405	24%	**.67**
Helvécia	58/374	16%	**.47**
Cinzento	117/927	13%	**.43**

A comunidade de Cinzento apresenta uma história um pouco diferenciada da comunidade de Helvécia e se assemelha à de Rio de Contas. Cinzento foi fundada por negros fugidos que se lançaram por um terra desconhecida, de difícil acesso, com poucas oportunidades de bem-estar, distante do centro comercial, carente de infraestrutura. A endogamia foi utilizada como mecanismo de manutenção da posse da terra e princípio básico de coesão do grupo. Além disso, a energia elétrica chegou tardiamente em Cinzento (1999). Por outro lado, a comunidade de Rio de Contas dispõe de uma estrutura turística, graças às belezas naturais da região. As comunidades afro-brasileiras ali localizadas apresentam uma história de isolamento, mas não tão acentuada como a comunidade de Cinzento. O turismo fez com que os falantes das comunidades de Barra e Bananal entrassem em contato com um maior contingente de falantes das normas urbanas, favorecendo a aplicação da regra de concordância. A comunidade de falantes de Rio de Contas estaria, portanto, mais sujeita à influência externa do que a comunidade de Cinzento, que, antigamente, via na presença do forasteiro uma ameaça à paz estabelecida na região.

A comunidade de Helvécia, atualmente, vem passando por transformações relativas ao deslocamento da atividade de subsistência para o plantio do eucalipto. A mudança econômica seria um fator de modificação da estrutura social e, portanto, linguística. O contato com padrões linguísticos prestigiados poderia influenciar a mudança da concordância segundo os padrões da língua-alvo.

Todo esse quadro só vem a corroborar a visão de uma realidade linguística brasileira polarizada, segundo a qual as variedades populares do português do Brasil, particularmente nas comunidades rurais afro-brasileiras, se caracterizam pela redução do seu sistema flexional, determinada, em sua origem, pelo contato entre línguas; em contraste com uma leve variação no uso das regras de concordância que se observa na norma urbana culta, de modo que a falta de concordância é fortemente estigmatizada nos ambientes urbanos cultos. Em função desse quadro, a difusão de padrões linguísticos e culturais a partir dos grandes centros urbanos sobre as variedades populares da cidade e do campo promove um processo de mudança em direção à aquisição das marcas de flexão de acordo com os padrões da língua culta.

Nesse contexto, os resultados da variável **sexo** revelaram que os homens e as mulheres possuem papéis diferentes no cenário da vida rural. Como é comum acontecer na sociedade brasileira, a mulher assume uma posição doméstica, representando o elo entre o marido e os filhos. Quando solteira, a liberdade da mulher é mais restrita do que a do jovem solteiro. Isso produz um ciclo de contato menos amplo, propiciando maior influência do meio familiar. Quando casada, exerce uma dupla jornada, uma na roça e outra em casa, espaços legítimos da presença feminina. Além disso, são elas mais propensas ao contato religioso, às devoções e à preservação do culto. Todas essas características fazem, certamente, com que o comportamento feminino seja mais conservador do que o masculino. Esse conservadorismo manifesta-se até nos costumes e no tratamento com as pessoas. As mulheres de Cinzento, sem exceção, usam lenços na cabeça, um resquício dos hábitos religiosos africanos, mas não demonstram ter consciência dessa prática, apenas usam porque aprenderam com seus antepassados. No momento das entrevistas, as mulheres mostraram-se menos receptivas a conversas que não girassem em torno do trato doméstico. O homem cinzentense, por outro lado, mantém uma relação mais próxima com outros grupos, indo à cidade, à feira, responsabilizando-se pelo comércio e, portanto, em maior contato com falantes de outras normas. Os dados obtidos nas comunidades estudadas confirmam essa hipótese, como se pode ver na Tabela 17.

Tabela 17 - Aplicação da regra de concordância verbal em comunidades rurais afro-brasileiras isoladas segundo a variável sexo

SEXO	Nº DE OC. / TOTAL	FREQ.	P.R.
Homens	153/806	19%	.56
Mulheres	120/900	13%	.45

Dessa forma, observa-se um cenário bem distinto daquele dos grandes centros urbanos dos países industrializados, em que as mulheres lideram a mudança em direção à norma de prestígio (CHAMBERS; TRUDGILL, 1980, p.97-98), corroborando a visão de que o papel da mulher em relação à mudança linguística deve ser considerado em cada situação cultural e sócio-histórica específica.[13]

14.2.3 Conclusão

Os resultados empíricos obtidos nesta análise contrariam frontalmente o quadro de perda da concordância apontado por Naro (1981) e por Naro e Scherre (2007) e revelam uma tendência para a aquisição da regra de concordância, em consonância com os achados de Guy (1981a), Vieira (1997) e Lucchesi (2000a), entre outros que têm realizado análises variacionistas de variedades populares do português brasileiro. Assim, evidências históricas e linguísticas apontam para um processo de mudança em direção à aquisição da regra de concordância nessas variedades, e não um processo contínuo e gradual de perda das marcas de concordância previsto pela deriva secular.

Por outro lado, o quadro de variação em relação à concordância verbal junto à 3ª pessoa do plural ajusta-se perfeitamente à visão da polarização do PB. A variação nas comunidades rurais afro-brasileiras isoladas enquadra-se na variação nas normas populares do PB, e é quantitativa e qualitativamente distinta da observada nas normas urbanas culta e semiculta, nas quais a variação é quantitativamente muito restrita, de modo que a falta de concordância é estigmatizada, e o quadro observado se afigura como o de uma variação estável, caracterizado por uma variação geracional que tende a se reproduzir no devir histórico, fazendo com que o indivíduo vá ajustando o seu comportamento linguístico ao padrão normativo, ao longo da sua vida. Já no polo popular, observa-se uma aquisição da regra segundo os padrões do português culto. Estruturalmente, esse processo estaria condicionado pelo princípio da saliência fônica, da transparência semântica e da coesão estrutural. No plano social, a influência dos centros urbanos, bem como a melhoria na infraestrutura do campo, tem uma

---

[13] Cf. Lucchesi, 2004c, p.192: "Generalizações do tipo '*as mulheres são mais inovadoras que os homens*' ou '*as mudanças linguísticas se propagam a partir de um ponto intermediário na escala social*' têm, a meu ver, um valor heurístico bastante questionável, pois a ação de um determinado fator social sobre um processo particular de mudança é determinada pela maneira específica através da qual esse fator se integra no conjunto complexo de interações que constitui o processo social em que a mudança acontece. [...] Portanto, o papel da mulher só pode de fato ser considerado dentro de cada realidade sociocultural específica e para cada caso particular de mudança. Qualquer generalização sobre o papel da mulher na mudança linguística em geral é extremamente perigosa para o entendimento da questão como processo histórico".

influência decisiva no curso da mudança. Essa tendência, já observada por Amaral (1920), que via a obsolescência do dialeto caipira em função da urbanização de São Paulo, poderá ser verificada nas comunidades rurais que estão cada vez mais importando os padrões e modelos urbanos. Os moradores da zona rural apresentam, via de regra, avaliações positivas em relação ao comportamento dos moradores da cidade quando se trata de educação e bem-estar.

Todo esse processo está alterando as feições do português popular do interior do país e, particularmente, o português afro-brasileiro (conforme definido no capítulo 2 deste livro), que tem sua origem ligada ao contingente de africanos que adquiriram o português de forma defectiva, produzindo uma variedade da língua portuguesa marcada pela redução na morfologia flexional do verbo. Tal variedade se teria fixado nas gerações seguintes até sofrerem os influxos atuais decorrentes da diminuição do isolamento em que se encontravam essas comunidades de fala. Portanto, os dados empíricos se articulam de forma bastante consistente com uma visão sócio-histórica do português brasileiro, que integra de maneira decisiva a diversidade étnica da população brasileira.

14.3 A concordância verbal com a primeira pessoa do plural

No que concerne à concordância com a 1ª pessoa do plural quando esta se atualiza na forma do pronome canônico *nós* ou de uma categoria vazia ligada a esse pronome anaforicamente, observa-se, no português do Brasil, um quadro semelhante ao que se observa com a concordância verbal junto à 3ª pessoa do plural: uma polarização sociolinguística, calcada na alta frequência de aplicação da regra de concordância na norma culta, em contraste com uma aplicação muito baixa da regra na norma popular, a que correspondem sistemas distintos de avaliação subjetiva das variantes e tendências de mudança igualmente diferenciadas em cada uma das normas.

A ausência de estudos sobre a variação na concordância verbal junto à 1ª pessoa do plural em materiais do Projeto NURC, por si só, já é um indicador de que a ocorrência desse fenômeno na norma culta seria extremamente lateral, ou mesmo irrelevante em termos estruturais. Analisando o comportamento linguístico de falantes com grau de escolaridade fundamental e médio de Porto Alegre e Panambi, no Rio Grande do Sul, Zilles, Maya e Silva (2000) encontraram uma frequência de 87% de utilização da desinência *-mos* (e de seu alomorfe *-mo*) com o sujeito representado pelo pronome *nós*. Note-se, em primeiro lugar, que essa frequência é superior à da aplicação da regra de concordância com a 3ª pessoa do

plural entre falantes com os mesmos níveis de escolaridade das cidades de Florianópolis e do Rio de Janeiro — 79% e 73%, respectivamente (MONGUILHOTT; COELHO, 2002; SCHERRE; NARO, 1997). Além disso, considerando apenas os falantes com ensino médio do estudo de Zilles, Maya e Silva (2000, p.209), o nível de aplicação da regra chega a 95%. Esses resultados, portanto, confirmam a hipótese de que a frequência de uso da regra de concordância verbal com a 1ª pessoa do plural entre os falantes escolarizados dos centros urbanos brasileiros é ainda maior do que a frequência de uso da regra de concordância junto à 3ª pessoa do plural. Isso provavelmente se deve ao fato de que o morfema de 1ª pessoa do plural *-mos* tem mais consistência fônica do que a marca de 3ª pessoa do plural, que, na maioria dos casos, se restringe à nasalização da sílaba átona final.

Mas o cenário de um elevado grau de aplicação da regra de concordância começa a mudar significativamente, quando se observa a periferia dos centros urbanos, com falantes com pouca ou nenhuma escolaridade. Ângela Rodrigues (1992) estudou o fenômeno em uma comunidade de favelados da periferia da cidade de São Paulo e encontrou um índice de aplicação da regra de 53% entre falantes analfabetos ou semianalfabetos. Bortoni-Ricardo (1985) encontrou praticamente a mesma frequência (56%) entre migrantes da zona rural que se estabeleceram na cidade satélite de Brazlândia, na periferia de Brasília. Esses falantes, com pouca ou nenhuma escolaridade, são definidos como *rurbanos*, dentro da classificação adotada pela autora.

A análise do fenômeno feita por Almeida (2005) em São Miguel dos Pretos, uma comunidade rural afro-brasileira do Rio Grande Sul, apresenta um elevado índice de aplicação da regra de concordância, na ordem de 73%.[14] Isso talvez possa ser explicado pelo fato de essa comunidade não ser representativa das comunidades rurais pobres do país, sobretudo das comunidades rurais afro-brasileiras isoladas do Norte e do Nordeste. Como admite a própria autora,

> [...] a alta frequência de concordância de São Miguel dos Pretos provavelmente se deve ao fato de que os moradores da comunidade [...] mantêm diferentes relações com pessoas de outras comunidades, tanto da zona rural quanto da zona urbana, devido à necessidade de emprego, de saúde, de educação, entre outras (ALMEIDA, 2005, p.114).

Na fala das comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia, aqui analisadas, a frequência de uso do morfema *-mos* (e alomorfes *-mo* e *-emo*) junto a sujeitos representados por *nós* é de apenas 18%. Esse índice parece

---

[14] A comunidade também exibe uma frequência de 81% de concordância com a 3ª pessoa do plural (ALMEIDA, 2005, p.136), que também foge ao padrão esperado.

ser mais representativo das comunidades rurais afro-brasileiras isoladas e das comunidades rurais pobres do interior do país.

Assim, configura-se também para a concordância verbal da 1ª pessoa do plural um quadro de polarização sociolinguística do Brasil, com os falantes urbanos com alto nível de escolaridade empregando a regra de concordância em um nível quase categórico, enquanto a variação no uso da regra na norma popular brasileira pode chegar a mais de 80%.

Um outro aspecto que diferencia a norma culta da norma popular brasileira é a frequência de emprego de alomorfes da forma padrão *-mos* da desinência de 1ª pessoa do plural. O apagamento da consoante constritiva da coda silábica (o *-s* final do morfema) e a alternância da vogal temática, em formas como *nós cheguemo* em vez de *nós chegamos*, são muito mais frequentes na fala popular do que na culta, até porque, entre os falantes urbanos escolarizados, essas formas são bastante estigmatizadas, sobretudo o alomorfe *-emo*, tido como símbolo da fala "caipira". Assim, no estudo de Zilles, Maya e Silva (2000, p.206), com falantes escolarizados de Porto Alegre e Panambi, 53% dos 87% do total de ocorrências de desinências verbais de 1ª pessoa do plural correspondem à forma padrão *-mos*, enquanto na comunidade rural afro-brasileira de São Miguel dos Pretos, no mesmo estado, só 26% do total de 79% de desinências eram da forma padrão (ALMEIDA, 2005, p.91).

Outra diferença entre a norma culta e o português popular brasileiro é que a variação na concordância verbal com a 1ª pessoa do plural se dá tanto com a expressão *a gente*, quanto com o pronome canônico *nós*, como exemplificado abaixo:

(21) As idéa, [*nós*] num *combinamos*. (RC-05)[15]

(22) Não, *nós* aqui não *dançava*, num existia dança. (CZ-11)

(23) Ah, meu pai, meu pai trabalhava quase igualmente *a gente trabaia*. (SP-12)

(24) Se 'doece um tomém, *a gente tamo* lá. (HV-04)

Zilles, Maya e Silva (2000, p.201) informam que "nos dados de Porto Alegre do VARSUL (24 entrevistas analisadas), há apenas uma ocorrência de sujeito *a gente* combinada com a desinência *-mos*, dita jocosamente pelo informante". Portanto, na norma culta, tal combinação está praticamente ausente, até porque é muito estigmatizada, ao passo que, na fala popular, ela pode ocorrer, mas não é muito frequente (Bortoni-Ricardo, 1985, p.211-212), não por ser avaliada negativamente por seus utentes, mas porque a norma dos falantes com pouca ou nenhuma escolaridade se caracteriza exatamente pelo reduzido uso de desinências verbais de número e pessoa.

---

[15] Os exemplos são extraídos dos dados da amostra de fala recolhida junto às comunidades analisadas e são identificados pela sigla das comunidades (CZ, para Cinzento; HV, para Helvécia; RC, para Rio de Contas; e SP, para Sapé) seguida do número do informante. Uma apresentação do *corpus* é feita no capítulo 5 deste livro.

Assim, até para poder traçar paralelos com outros estudos, só serão apresentados aqui os resultados da variação com a concordância verbal junto ao pronome canônico *nós*, realizado ou não, estruturando-se a variável dependente da seguinte maneira:

1. com marcação morfológica

(25) *Nós so**mos*** seis irmão. (CZ-06)

(26) *Nós fo**mo*** tudo em Helvécia, mas já faz muito tempo… (HV-19)

(27) Aí, *comece**mo*** a gostá e *case**mo***. (CZ-03)

2. sem marcação morfológica

(28) porque *nós* só ***vai*** sabê quanto vai saí pa cada produtô quando ***fô*** assiná o pojeto. (CZ-06)

A análise dessa variável, que segue os pressupostos da Sociolinguística Quantitativa (WEINREICH; LABOV; HERZOG , 2006 [1968]; LABOV, 1972a, 1982, 1994; SANKOFF, 1988), tem como *corpus* uma amostra de fala constituída por 24 entrevistas de tipo sociolinguístico realizadas com moradores com pouca ou nenhuma escolaridade de quatro comunidades rurais afro-brasileiras isoladas de quatro diferentes regiões do Estado da Bahia, no período de 1992 a 2002. Na composição da amostra, os informantes foram divididos equitativamente por sexo e faixa etária. Desse modo, a amostra de fala foi estruturada de acordo com as seguintes variáveis: (i) comunidade: Sapé (Município de Valença, no Recôncavo Baiano), Helvécia (Município de Nova Viçosa, no Extremo Sul), Barra e Bananal (Município de Rio de Contas, na Chapada Diamantina), Cinzento (Município de Planalto, na região do Semiárido); (ii) sexo: masculino e feminino; (iii) idade: faixa 1, 20 a 40 anos; faixa 2, 41 a 60 anos; faixa 3, mais de 60 anos; (iv) escolaridade: semianalfabeto e analfabeto; (v) estada fora da comunidade: para figurar com valor positivo, o falante deveria ter vivido pelo menos seis meses fora da comunidade.

Na amostra de fala analisada, foram depreendidas 480 ocorrências de formas verbais relacionadas com um sujeito representado pelo pronome *nós*, realizado ou não foneticamente. Considerando como ocorrências da aplicação da regra de concordância tanto a presença do morfema canônico *-mos*, quanto os alofones não padrão *-mo* (com supressão do *-s* final), quanto *-emo* (com alteração da vogal temática nos verbos da 1ª conjugação), os resultados da variável dependente foram os constantes na Tabela 18a.

Tabela 18a - Frequência de uso do morfema verbal de primeira pessoa do plural no português afro-brasileiro

MORFEMA	Nº DE OC. / TOTAL	FREQUÊNCIA
-mos ~ -mo ~ -emo	84/480	**18%**
∅	396/480	**82%**

Como já foi dito, a frequência de aplicação da regra de concordância é bastante baixa, encontrando-se praticamente no mesmo nível de concordância com a 3ª pessoa do plural nessas mesmas comunidades, que ficou em 16% (cf. seção anterior). Vale ressaltar ainda que, desmembrando-se os alofones do morfema de 1ª pessoa do plural, o morfema padrão corresponde a apenas 2% do total de ocorrências, o mesmo percentual do alomorfe *-emo*. Assim, é muito frequente na fala das comunidades analisadas o apagamento do *-s* final do morfema, que caracteriza a fala popular brasileira, como se pode constatar na Tabela 18b.

Tabela 18b - Frequência de uso dos alomorfes verbais de primeira pessoa do plural no português afro-brasileiro

MORFEMA	Nº DE OC. / TOTAL	FREQUÊNCIA
-mo	64/480	**13%**
-mos	10/480	**2%**
-emo	10/480	**2%**
∅	396/480	**82%**

Considerando-se o total de 18%[16] de aplicação da regra de concordância, serão apresentados a seguir os resultados das variáveis linguísticas e sociais que afetam a aplicação da regra nas comunidades de fala analisadas.

14.3.1 O encaixamento linguístico

Na análise do encaixamento linguístico da variação na concordância verbal com a 1ª pessoa do plural no português afro-brasileiro, o programa de regras variáveis VARBRUL selecionou como relevantes as seguintes variáveis explanatórias: (i) realização e posição do pronome sujeito, (ii) paralelismo discursivo; e (iii) saliência fônica. Os resultados da análise quantitativa de cada uma dessas variáveis e a sua interpretação são apresentados nas próximas subseções.

14.3.1.1 Realização e posição do pronome sujeito

A realização e a posição do sujeito é uma variável presente em qualquer análise da aplicação da regra de concordância verbal. Nesse caso, dois princípios antagônicos estão atuando. De um lado, a presença do sujeito imediatamente antes do verbo favorece a aplicação da regra de concordância, em função da facilidade que opera nesses casos no processamento linguístico do falante. De outro lado, a não realização fonética do sujeito também favoreceria a realização da concordância,

[16] O total chega a dezoito por cento em função das frações desprezadas nos valores aproximados do quadro anterior.

tendo-se em conta uma lógica funcional, pois o morfema de número e pessoa supriria a ausência de informação com o sujeito nulo. Já a posposição do sujeito, que tem se mostrado um fato bastante desfavorecedor da concordância verbal, ocorreu em um nível insignificante na amostra analisada. A posposição do sujeito é mais frequente quando um SN desempenha esta função sintática, sendo essa inversão mais rara no caso do sujeito pronominal. Isso pode explicar a baixíssima ocorrência da posposição na base de dados desta análise, na qual foram considerados apenas os seguintes fatores:

(1) pronome sujeito realizado imediatamente antes do verbo

(29) É aonde *nós tamos* por aí até o dia que Deus querê. (RC-24)

(2) pronome sujeito separado do verbo por algum constituinte[17]

(30) Não, *nós* ***aqui não*** *dançava*. Num existia dança. (CZ-11)

(3) sujeito nulo

(31) Quano prantô em novembo, nós *cói* em feverêro, *cói* em janêro. (CZ-11)

Os resultados da análise quantitativa são apresentados na Tabela 19.

Tabela 19 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável realização e posição do sujeito

REALIZAÇÃO E POSIÇÃO DO SUJEITO	Nº DE OC. / TOTAL	FREQ.	P.R.
Sujeito nulo	38/99	38%	**.88**
Sujeito imediatamente antes do verbo	38/304	13%	**.38**
Sujeito antes do verbo (não adjacente)	08/72	11%	**.37**
Total	**84/475**	**18%**	—

Nível de significância: .009.

Na amostra de fala analisada prevaleceu a lógica funcional, sendo o contexto de sujeito nulo aquele em que a regra de concordância verbal mais é aplicada, correspondendo a mais que o dobro da frequência geral (38% *versus* 18%; com peso relativo de .88). A presença do pronome sujeito, imediatamente antes do verbo ou separado por algum constituinte, acabou por desfavorecer a aplicação da regra de concordância, com 13% e 11% de frequência e peso relativo de .38 e .37, respectivamente.

Esses resultados são análogos aos encontrados em outros estudos sobre o fenômeno no PB. Bortoni-Ricardo (1985) encontrou 84% de concordância com o sujeito nulo, contra apenas 47% de pronome *nós* anteposto ao verbo. No estudo

[17] Foram aí incluídas as pouquíssimas ocorrências em que o pronome *nós* é retomado por um pronome relativo, codificadas inicialmente como um fator específico, mas depois amalgamadas em função da sua baixa ocorrência e da semelhança com o contexto da interposição de um constituinte entre o pronome e o verbo.

de Ângela Rodrigues (1992), também o sujeito nulo foi o fator que mais favoreceu a concordância, com 74% de frequência e peso relativo de .81, contra 47% de frequência e peso relativo de .45 para o *nós* explícito. Almeida (2005) também observa a prevalência do mesmo princípio com a frequência de concordância de 83% para o sujeito nulo, contra 63% para a presença do pronome *nós* anteposto ao verbo.

Portanto, constata-se a prevalência do princípio funcional em que a ausência do pronome reforça a necessidade da indicação de pessoa e número do sujeito através da desinência verbal. Esses resultados contradizem as conclusões de Scherre e Naro (1993), com base em seus estudos da concordância verbal com a 3ª pessoa do plural, de que o "princípio do paralelismo formal", segundo o qual "marcas levam a marcas e zeros levam a zeros", suplanta o "princípio da economia" na "repetição de marcas formais"; do mesmo modo que desautorizam a sua generalização de uma "tendência em direção a marcas redundantes" no PB (NARO; SCHERRE, 2007, p.177). Não obstante, a ação do paralelismo discursivo não deixou de ser relevante, como se pode ver na próxima subseção.

14.3.1.2 Paralelismo discursivo

O princípio do paralelismo discursivo, segundo o qual o falante tende a repetir suas escolhas na cadeia da fala, tem-se mostrado relevante na análise variacionista da concordância verbal. Em uma cadeia de orações com referência ao mesmo sujeito, se o falante utiliza o morfema verbal de número e pessoa na primeira oração, ele tende a utilizar esse morfema nas orações seguintes, de acordo com o princípio proposto por Scherre e Naro (1993) de que "marcas levam a marcas, e zeros levam a zeros". Com base nesse princípio, essa variável explanatória foi estruturada da seguinte maneira:

1. forma verbal com morfema de 1ª pessoa do plural na oração anterior

(32) Aí *nós entramo* pra dentro, aí *nós ficamo* lá, *nós ficamo* lá, aí a pouco ININT aquela... aquela bagunça... é... aqueles meninado... é home com mulé... (HV-07)

2. forma verbal sem morfema de 1ª pessoa do plural na oração anterior

(33) *nós saiu* daqui, *foi* assisti jogo no Inácio, e hoje *nós asseste* aqui. (CZ-03)

Os resultados quantitativos apresentados na Tabela 20 confirmam o princípio do paralelismo discursivo, já que a presença de uma forma verbal marcada na oração anterior foi o fator que mais favoreceu a aplicação da regra de concordância verbal, com uma frequência de 48% (e peso relativo de .72), frente a uma frequência geral de 18%. Deve-se destacar também que, por fatores pragmáticos e funcionais, o morfema de pessoa e número também é bastante usado na primeira referência (com frequência de 19% e peso relativo de .62), já que, na primeira menção do sujeito, as marcas de referência tornam-se mais necessárias.

Tabela 20 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável paralelismo discursivo

PARALELISMO DISCURSIVO	Nº DE OC. / TOTAL	FREQ.	P.R.
Forma verbal marcada na oração anterior	13/27	48%	**.72**
Primeira referência	52/276	19%	**.62**
Forma verbal não marcada na oração anterior	19/177	11%	**.28**
Total	**84/480**	**18%**	—

Nível de significância: .009.

14.3.1.3 Saliência fônica

A saliência fônica, proposta por Lemle e Naro (1977) para o português, pode ser considerada hoje uma variável clássica nas análises variacionistas dos fenômenos de concordância. Segundo esse princípio, o falante tende a fazer mais a concordância com aqueles itens lexicais que se flexionam de forma mais acentuada, do ponto de vista fônico, conjugando-se a posição do acento e a substância fônica do morfema flexional. Para a flexão verbal de 1ª pessoa do singular, foi adotada a seguinte escala, proposta por Naro, Görski e Fernandes (1999, p.203):

(1) a oposição V/V*mos* não é tônica nas duas formas

(Ex.: *falava/falávamos*; *fosse/fôssemos*; *ia/íamos*)

(2) a oposição V/Vmos é tônica em uma das formas

(Ex.: *fala/falamos*; *trouxe/trouxemos*; *assiste/assistimos*; *fique/fiquemos*)

(3) a oposição V/V*mos* é tônica nas duas formas

(Ex.: *está/estamos*; *tem/temos*; *fazer/fazermos*; *faz/fazemos*; *será/seremos*; *dá/damos*; *vê/vemos*)

(4) a oposição V/V*mos* é tônica nas duas formas, e na primeira ocorre uma semivogal que não ocorre na segunda

(Ex.: *comeu/comemos*; *partiu/partimos*; *vai/vamos*; *foi/fomos*)

(5) a oposição V/V*mos* é tônica nas duas formas, e também ocorre mudança no radical

(Ex.: *veio/viemos*; *é/somos*; *fez/fizemos*)

Os resultados da Tabela 21 exibem uma perfeita sintonia com a escala proposta, o que é notável, já que nem sempre essa variável exibe um comportamento escalar, conforme constatado por Naro e Scherre (1999), apresentando por vezes um resultado contrastivo entre dois extremos. Deve-se destacar ainda que a frequência de aplicação da regra aumenta bastante a partir do nível 3 de saliência, em que a oposição se dá na sílaba tônica das duas formas em contraste; a partir daí, a frequência da concordância passa de 5% para a casa dos 20%, e o peso relativo salta de .25 para mais de 70 pontos.

Tabela 21 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável saliência fônica

SALIÊNCIA FÔNICA	Nº DE OC. / TOTAL	FREQ.	P.R.
1. falava/falávamos	01/96	01%	**.09**
2. fala/falamos	04/87	05%	**.25**
3. está/estamos	22/86	26%	**.71**
4. comeu/comemos	41/163	25%	**.74**
5. veio/viemos; é/somos	16/48	33%	**.80**
Total	**84/480**	**17%**	—

Nível de significância: .009.

14.3.2 O encaixamento social

Na análise do encaixamento do processo de variação na concordância verbal junto à 1ª pessoa do plural nas comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia, os resultados de duas variáveis sociais podem ser identificados como indícios de um processo de mudança em curso no sentido do incremento da aplicação da regra de concordância. Entretanto, a força com que o processo de substituição de *nós* por *a gente* se vem implementando nessas comunidades (cf. capítulo 19) pode vir a sustar esse processo ainda incipiente de mudança.

Para se chegar a um diagnóstico sobre um processo de variação sincrônica, nos termos da dicotomia *variação estável versus mudança em curso*, no tipo de abordagem denominada de *tempo aparente*, o ponto de partida é a distribuição geracional das frequências de uso das variantes na comunidade (LABOV, 1981; LUCCHESI, 2001b; entre outros).

Os resultados da variável **faixa etária** nas comunidades analisadas revelam uma ligeira inclinação no sentido do aumento do uso da regra da concordância verbal, na medida em que se passa dos falantes mais velhos para os falantes mais jovens, o que poderia ser considerado um indício de um processo de implementação da regra de concordância na gramática das comunidades de fala analisadas. Entretanto, a diferença percentual entre as faixas etárias é mínima, como se pode ver na Tabela 22, e a variável não foi selecionada pelo programa de regras variáveis VARBRUL como estatisticamente significativa, no processamento quantitativo dos dados.

Tabela 22 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável faixa etária

FAIXA ETÁRIA	Nº DE OC. / TOTAL	FREQ.
20 a 40 anos	20/104	19%
41 a 60 anos	34/188	18%
Mais de 60 anos	30/188	16%
Total	**84/480**	**17%**

Na Figura 3, pode-se visualizar a ligeira inclinação em favor do incremento do uso da regra de concordância nas comunidades analisadas, mas é necessário recorrer às outras variáveis sociais para se buscar um diagnóstico mais consistente.

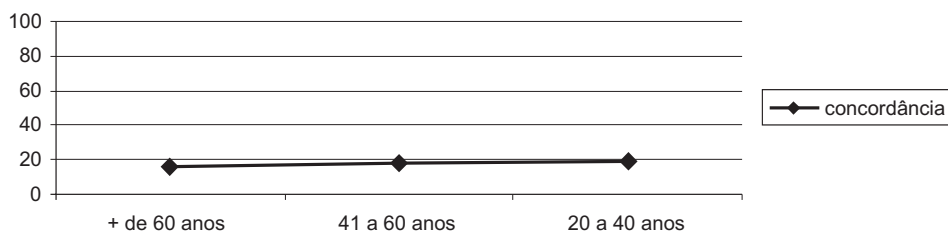



Figura 3 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável faixa etária do falante (em percentuais)

Os resultados quantitativos da variável **sexo** revelam que os homens tendem a empregar mais a regra de concordância verbal que as mulheres nas comunidades rurais afro-brasileiras estudadas, apesar de essa variável também não estar no nível máximo de tolerância definido pelo logaritmo da verossimilhança (nível de significância de até .050), como se pode ver na Tabela 23.

Tabela 23 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável sexo do falante

SEXO	Nº DE OC. / TOTAL	FREQ.	P.R.
Homens	51/198	26%	**.56**
Mulheres	33/282	12%	**.45**
Total	**84/480**	**17%**	—

Nível de significância: .176.

Os estudos já desenvolvidos sobre as comunidades rurais revelam que os homens tendem a liderar as mudanças em favor do padrão normativo, pois mantêm mais contatos fora da comunidade do que as mulheres. São os homens que se deslocam mais para as feiras para comercializar os eventuais excedentes de sua produção agrícola familiar e se dirigem para os grandes centros urbanos em busca de trabalho. Já as mulheres tendem a ficar mais circunscritas ao universo doméstico e ao trabalho na roça. Portanto, os homens acabam por sofrer mais influências externas, liderando o processo de assimilação dos padrões linguísticos urbanos.

Outra evidência forte de que o quadro seria de uma mudança aquisicional de cima para baixo, a partir da crescente influência dos grandes centros urbanos, é obtida com os resultados da variável estada fora da comunidade. Através dessa variável, que foi selecionada pelo VARBRUL como estatisticamente relevante (com nível de significância de .009), pode-se constatar que aqueles indivíduos que viveram fora das comunidades estudadas por um período superior a seis meses (geralmente

trabalhando no Sul do país) são os que mais empregam a regra de concordância verbal, como se pode ver na Tabela 24. Essa seria então a evidência mais forte em favor de uma mudança de fora para dentro da comunidade em função de um processo de nivelamento linguístico que se expande a partir dos grandes centros urbanos.

Tabela 24 - A concordância verbal com a 1ª pessoa do plural no português afro-brasileiro segundo a variável estada fora da comunidade

ESTADA FORA DA COMUNIDADE	Nº DE OC. / TOTAL	FREQ.	P.R.
Pelo menos por seis meses	55/244	23%	**.59**
Não	29/236	12%	**.40**
Total	**84/480**	**17%**	—

Nível de significância: .009.

14.3.3 Conclusão

Apesar das evidências em favor de um processo de incremento de emprego do morfema verbal da 1ª pessoa do plural nas comunidades rurais afro-brasileiras isoladas, os resultados quantitativos não fornecem um quadro muito nítido. A razão disso pode ser a interferência do vigoroso processo de substituição do pronome *nós* pela expressão nominal *a gente*, gramaticalizada na função de pronome pessoal. Como se pode ver no capítulo 19 deste livro, o uso de *nós*, na fala dos indivíduos entre 20 e 40 anos das comunidades analisadas, restringe-se a pouco mais de 10% do total de ocorrências, predominando largamente o uso da forma inovadora *a gente*, que se conjuga normalmente com a forma não marcada da 3ª pessoa do singular. E seriam exatamente esses indivíduos que deveriam liderar o processo de mudança em favor do emprego da regra de concordância. Como eles podem adquirir uma regra de concordância que está relacionada com um pronome que está caindo em desuso na sua fala? Isso pode explicar por que o incremento da regra de concordância verbal com a 1ª pessoa é tão incipiente nessas comunidades, diferentemente do que já se observou para o emprego dos morfemas verbais da 1ª pessoa do singular e da 3ª pessoa do plural, em que a curva ascendente indicando a mudança aquisicional é clara (cf. seções 14.1 e 14.2, acima).

Entretanto, os estudos de outras variedades do PB apontam para um quadro mais claro de aquisição da regra de concordância com a 1ª pessoa do plural. Na análise de Bortoni-Ricardo (1985), na cidade satélite de Brazlândia-DF os jovens apresentam uma frequência de concordância de 82% contra 48% dos adultos; os homens, mais inseridos no mercado de trabalho que as mulheres, que ficam mais restritas ao ambiente doméstico, lideram a mudança em direção à variante padrão, com 66% de concordância contra 42% das mulheres. Na periferia da cidade de São Paulo, Ângela Rodrigues (1992, p.169) também observou que os homens fazem

mais a concordância com a 1ª pessoa do plural que as mulheres (com 63% de frequência e peso relativo de .57), já que os homens têm "mais opções profissionais na cidade grande que as mulheres". Para a comunidade rural afro-brasileira do Rio Grande do Sul, Almeida (2005, p.118-119) chega à conclusão de que

> [...] diante da trajetória histórica e social da comunidade e da diferença dos resultados entre a geração mais velha e a geração mais nova, jovens e adultos, podemos observar que a variedade falada em São Miguel passa pelo processo de aquisição da concordância verbal, ainda que a variável faixa etária não tenha sido selecionada pelo programa de cálculo estatístico.[18]

Considerando todos os resultados desse conjunto de estudos sobre a concordância verbal com a 1ª pessoa do plural, pode-se concluir que, na norma popular, existe uma tendência ao incremento do uso da desinência verbal de pessoa e número com os sujeitos representados pelo pronome *nós*. Entretanto, esse processo de mudança pode estar sofrendo uma forte sangria, em função da intensidade da substituição do pronome *nós* pela expressão *a gente*, sobretudo na fala dos mais jovens. Esse quadro de uma tendência de mudança aquisicional da regra de concordância verbal, que se observa na norma popular do PB, contrasta com o uso praticamente categórico da regra de concordância junto ao sujeito representado pelo pronome *nós* na norma culta. Portanto, a polarização sociolinguística do Brasil, constatada em relação à regra de concordância verbal junto à 3ª pessoa do plural, também se confirma em relação à 1ª pessoa do plural.

Da mesma forma, o baixíssimo uso do morfema de número e pessoa ajusta-se perfeitamente à interpretação sócio-histórica de que, na formação dos padrões linguísticos das comunidades rurais afro-brasileiras isoladas, a morfologia verbal de pessoa e número do português foi fortemente afetada pela erosão que ocorre normalmente nas situações de contato entre línguas massivo, radical e abrupto. Tal processo não teria levado à eliminação dessa morfologia, como ocorre nos casos típicos de crioulização, mas teria produzido um quadro de ampla variação, que caracteriza os processos de transmissão linguística irregular de tipo leve (cf. capítulo 3 deste livro). Esse quadro de ampla variação estaria sendo mitigado atualmente em função do nivelamento linguístico em curso, provocado pela ampla difusão dos padrões urbanos para todas as regiões do país.

---

[18] A autora encontrou uma frequência de 66% no emprego do morfema de 1ª pessoa do plural na fala da geração mais velha, contra 79% e 77% na fala de adultos e jovens, respectivamente.

15 A concordância em estruturas passivas e de predicativo do sujeito

Dante Lucchesi

Este capítulo trata da variação na concordância de número e gênero em predicativos do sujeito e estruturas passivas, tendo como universo de observação quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia. Na análise, observam-se diferenças nas frequências de uso dessas regras de concordância em diferentes variedades do português brasileiro decorrentes da clivagem social que marca o país. Por outro lado, são observados significativos paralelos estruturais entre a aplicação das regras de concordância no interior do Sintagma Nominal (SN) e nas estruturas de predicação aqui analisadas. Desse modo, este capítulo estrutura-se da seguinte forma: numa primeira seção, são referidos os efeitos do contato entre línguas sobre a aplicação das regras de concordância nominal no português brasileiro, considerando-se a sua polarização sociolinguística; na segunda seção, introduz-se o tópico da concordância nominal em estruturas passivas e de predicativo do sujeito no português brasileiro; em seguida, na seção 15.3, são apresentados os resultados da análise variacionista no universo das comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia, compreendendo os resultados da frequência da aplicação da regra de concordância de número — bem inferiores aos constatados em análises do comportamento de falantes urbanos escolarizados — e os resultados da variação na concordância de gênero — fenômeno que não se observa nas cidades brasileiras, em um nível significativo; na conclusão, faz-se uma síntese e uma interpretação de conjunto dos resultados apresentados nesta análise, em cotejo com outras análises sociolinguísticas sobre a variação da concordância nominal no português brasileiro.

15.1 O contato entre línguas e a variação na concordância nominal no português brasileiro

Dentre os níveis mais afetados na estrutura gramatical do português brasileiro pela transmissão linguística irregular deve-se destacar a significativa

erosão provocada na morfologia flexional do nome e do verbo e *a fortiori* nas regras de concordância que se apoiam nesses morfemas (cf. capítulo 3 deste livro). O amplo quadro de variação na concordância nominal e verbal decorrente desse processo de transmissão linguística irregular de tipo leve constitui uma verdadeira linha demarcatória na realidade linguística brasileira, dividindo-a em duas grandes normas: a **norma culta**, só indiretamente afetada por esse processo, e a **norma popular**, cuja formação é indelevelmente marcada por processos de mudança induzidos pelo contato entre línguas (cf. capítulo 1 deste livro).

A regra de concordância de número no interior do SN oferece um quadro bastante significativo a esse respeito. Enquanto a análise de Scherre (1988), sobre os materiais do PEUL,[1] revelou que falantes urbanos com escolaridade fundamental e média da cidade do Rio de Janeiro fazem a concordância nominal numa proporção de aproximadamente 70% dos casos, a frequência de uso da regra de concordância nominal na comunidade afro-brasileira de Helvécia, no Extremo Sul da Bahia, não chega a 10% (cf. capítulo 11 deste livro).

No caso da concordância nominal, a categoria gramatical do número teria sido bem mais afetada pelo processo de transmissão linguística irregular do que a do gênero. Assim, enquanto a variação na concordância nominal de número atinge todas as variedades do português brasileiro, a variação na concordância de gênero é um fenômeno raro, circunscrito a algumas comunidades rurais isoladas. Mesmo entre as comunidades rurais afro-brasileiras, não é comum encontrar-se a variação na concordância de gênero no interior do SN, exemplificada em (1) abaixo, em níveis estruturalmente relevantes.

(1) *Muitas veze, 'duece um pessoa, num tem ambulança.*[2]

Em Helvécia, ainda é possível observar-se o ocaso de um processo de variação na concordância de gênero, que, apesar de exibir um índice de não aplicação da regra de apenas 5% no conjunto da comunidade, ainda guarda uma razoável sistematicidade (cf. capítulo 12 deste livro). Porém, uma outra investigação sobre uma amostra de fala das comunidades rurais afro-brasileiras do Município de Rio de Contas, no interior do Estado da Bahia, levantou 2.368 ocorrências de SNs femininos, dos quais, em apenas 27, a regra de concordância de gênero não era aplicada, o que corresponde a uma variação de apenas 1%. Esse nível de variação não é relevante em termos estruturais.

O fato de a concordância de número ter sido mais afetada do que a concordância de gênero pode ser atribuído ao caráter mais gramatical e redundante

[1] Programa de Estudo sobre o Uso da Língua, da Universidade Federal do Rio de Janeiro (OLIVEIRA; SCHERRE, 1996).

[2] Exemplo extraído da fala de membro da comunidade rural afro-brasileira isolada de Helvécia.

daquela. A flexão de número em português é mais gramatical do que a de gênero, no sentido de que aquela é mais geral do que esta. Diferentemente do que ocorre em relação ao gênero, praticamente todos os constituintes do SN se dobram à flexão de número. Para além dos determinantes e modificadores, quase todos os nomes e adjetivos se flexionam quanto ao número, ao passo que, para além dos determinantes (artigos, demonstrativos, possessivos, etc.), só os adjetivos de tema em *-o* e um conjunto reduzido de nomes se flexionam quanto ao gênero. Assim, a concordância de número torna-se mais redundante. Tomando-se como exemplo uma frase trivial em português como a apresentada em (2) abaixo, pode-se observar que, no SN *as coisas mais importantes*, o plural é marcado três vezes, enquanto a única marca morfológica do gênero feminino do SN é fornecida pelo artigo definido, já que o gênero feminino é um traço inerente ao item lexical *coisa*, não lhe sendo atualizado por qualquer morfema gramatical. Desse modo, a supressão da concordância de número no SN, apresentada em (3), não elimina qualquer informação do enunciado, seja ela referencial, seja ela gramatical, ao passo que a não realização da concordância de gênero, como expresso em (4), elimina a única marca morfológica que indica o gênero do SN, tornando a frase agramatical para a grande maioria dos falantes do português no Brasil.

(2) Pra mim, *as* coisa*s* mais importante*s* são o amor e a amizade.

(3) Pra mim, *as* coisa mais importante são o amor e a amizade.

(4) Pra mim, *os* coisa mais importante são o amor e a amizade.

Essa diferença no grau de variação entre as regras de concordância de número e gênero no plano dos constituintes do SN também se verifica quando a relação de concordância envolve um SN sujeito e um predicativo ou um particípio passado em estruturas passivas. Esse é apenas um dos paralelos interessantes que se podem traçar entre esses dois níveis sintagmáticos de aplicação das regras de concordância nominal.

15.2 A variação na concordância nominal em estruturas passivas e de predicativo do sujeito no português brasileiro

O nível de variação na aplicação da regra de concordância de número em estruturas passivas e de predicativo do sujeito, respectivamente exemplificadas em (5) e (6) abaixo, é maior do que o que se verifica no interior do SN.

(5) **Aquelas meninas** já foram **escolhidas** pela comissão.

(6) **Estes livros** são mais **caros**.

Trabalhando com base na amostra Censo do PEUL, que reúne entrevistas com 64 falantes da cidade do Rio de Janeiro dos dois sexos, de três diferentes faixas etárias e de nível de escolaridade de 1 a 11 anos, Scherre (1991) identificou um nível de variação da ordem de 50% na aplicação da regra de concordância de número em estruturas passivas e de predicativo do sujeito. Dias (1996) encontrou um nível semelhante de aplicação da regra em estruturas passivas e de predicativo em um *corpus* semelhante da Região Sul do país: 327 casos de concordância de número em um total de 768 ocorrências, o que corresponde a 43% do total. Desse modo, o nível de aplicação da regra nessas estruturas é inferior ao nível que se observa no interior do SN, pois, na mesma amostra, Scherre (1988) identificou uma frequência de 72% de aplicação da regra de concordância de número nesse nível da estrutura linguística. Essa diferença pode ser explicada por uma maior coesão estrutural que se observa no interior do SN, sobretudo no plano da relação Especificador-Núcleo, respectivamente ocupadas por um determinante (artigo, demonstrativo, etc.) e um nome.[3] Entretanto, quando o SN exibe uma estrutura do tipo Determinante+Nome+Modificador, o adjetivo que está em adjunção à direita do nome participa da relação de concordância no mesmo nível que um predicativo ou um particípio passado em uma estrutura passiva. Esse interessante paralelo estrutural foi observado nesta análise da concordância em estruturas desse tipo na fala de comunidades rurais isoladas no interior do Estado da Bahia.

15.3 A variação na concordância nominal em estruturas passivas e de predicativo do sujeito no português afro-brasileiro

Esta análise quantitativa da variação na concordância de número e de gênero em estruturas da voz passiva e de predicativo do sujeito tem como base empírica amostras de fala vernácula de três comunidades rurais afro-brasileiras isoladas de três diferentes microrregiões do interior do Estado da Bahia: Helvécia, na Zona da Mata, no litoral sul, ligada originalmente à cultura agroexportadora do café; Cinzento, na região do Semiárido, zona de pecuária; e Barra e Bananal, comunidades do Município de Rio de Contas, na Chapada Diamantina, antiga região de mineração. Essas comunidades vivem, nos dias de hoje, basicamente da agricultura de subsistência, particularmente da mandioca, da qual produzem artesanalmente a

[3] Análises variacionistas sobre a concordância no SN revelam subsidiariamente que entre 70% e 90% dos SNs produzidos na fala vernácula estão presentes apenas dois constituintes, na maioria dos casos, um determinante e um nome: *as meninas*, *aquela casa*, *uns problemas*, etc.

farinha, e a grande maioria dos seus membros descende diretamente de africanos trazidos para o Brasil como escravos. Há indícios de que as comunidades de Rio de Contas e Cinzento tenham a sua origem em antigos quilombos — agrupamentos de escravos foragidos. A comunidade de Helvécia se constituiu após o fim da antiga Colônia Leopoldina, formada na região por colonos alemães, suíços e franceses para o cultivo e exportação de café no século XIX.[4]

As amostras de fala são constituídas por entrevistas informais com doze informantes de cada comunidade, perfazendo um total de 36 entrevistas.[5] Os informantes são distribuídos equitativamente pelos dois sexos e três faixas etárias: de 20 a 40 anos; de 41 a 60 anos; e de mais de 60 anos. A análise foi feita com base na metodologia da Sociolinguística Variacionista (LABOV, 1972a, 1994), com recurso ao pacote de Programas VARBRUL, para o tratamento estatístico dos dados (NARO, 2003; SCHERRE; NARO, 2003).

15.3.1 A variação na concordância nominal de número em estruturas passivas e de predicativo do sujeito no português afro-brasileiro

No que concerne à variação na concordância de número com predicativos e em estruturas da voz passiva, a fala das comunidades rurais afro-brasileiras apresenta um nível de aplicação da regra bem inferior ao observado por Scherre (1991) e Dias (1996) com falantes urbanos escolarizados. Não se considerando os SNs que funcionam como predicativos, como exemplificado em (7), o índice de aplicação da regra é de apenas 1%, o que não é relevante em termos estruturais.[6]

(7) nós que somos *as pessoas moreno*, *as pessoa mais esquecido do Brasil*...

Assim sendo, pode-se dizer que a concordância de número com predicativos e particípios em estruturas da voz passiva ainda no final do século XX estava

---

[4] Uma descrição mais ampla dessas comunidades é feita no capítulo 5 deste livro.

[5] A metodologia empregada na constituição dessas amostras de fala é descrita no capítulo 5 deste livro.

[6] Na análise foram depreendidas 276 ocorrências de predicativos e particípios passivos ligados a um sujeito no plural. Nesse universo, apenas quatro ocorrências exibiram as marcas de concordância, conforme exemplificado em (7) a. e b. abaixo, o que corresponde a 1,14% do total.

(7) a. Ela enfeza com nós, que nós é muitos *traquinos*.

b. ... tem muitas pessoas que são *morenos*... e... é uma pessoa que tem muito valô.

virtualmente ausente da gramática das comunidades rurais afro-brasileiras.[7] Esse fato, comparado com a frequência média verificada em dialetos urbanos, reforça, por um lado, a polarização sociolinguística do Brasil proposta por Lucchesi (2001, 2002a); por outro lado, aponta para a relação histórica em que as variedades do português brasileiro ligadas mais diretamente às situações de contato linguístico na sua formação sofreram um processo de erosão mais intenso no uso das regras de concordância, um dos aspectos da gramática mais imediatamente afetados em situações de contato (cf. capítulo 3 deste livro).

Já no nível da estrutura linguística, verifica-se um interessante paralelo, quando se comparam os índices de concordância de número dos predicativos e particípios passivos com o grau de marcação do plural em adjetivos que se colocam em adjunção à direita do nome núcleo na estrutura do SN. Segundo Lobato (1994), esses elementos em adjunção à direita do nome seriam mais refratários à regra de concordância nominal, por estarem menos integrados à estrutura do SN. A Tabela 1, extraída da análise quantitativa que Andrade (2003) realizou sobre a variação na concordância de número no interior do SN na comunidade de fala de Helvécia, confirma esse princípio, revelando que a marcação do plural nos modificadores à direita do nome é praticamente nula na gramática da comunidade de fala.

Tabela 1 - Padrão de marcação do número no SN em Helvécia-BA

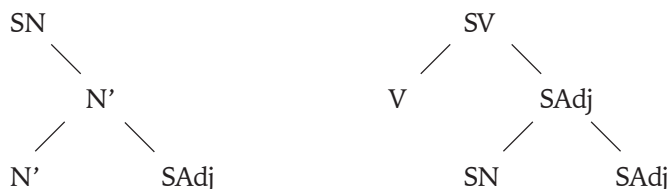
POSIÇÃO DO CONSTITUINTE	Nº DE OCORRÊNCIAS	%	PESO RELATIVO
Primeira posição antes do núcleo (*tem muitos ano...*)	1.154/1.196	96	.94
Segunda posição à esquerda do núcleo (*do meus irmão*)	25/37	68	.67
Núcleo em primeira posição (*mais cereais miúdo*)	3/8	38	.62
Núcleo em segunda posição (*essas festa mesmo de aruanda*)	116/1.434	8	.15
Núcleo em terceira posição em diante (*Aí só esses meus primo aí*)	10/115	9	.12
Modificadores à direita (*umas coisa velha*)	2/103	2	.00

Fonte: Andrade, 2003, p.94.

Note-se que o nível de marcação do plural nos adjetivos à direita do nome é o mesmo da concordância de número dos particípios passivos e predicativos do

[7] Como as análises sociolinguísticas dessas comunidades têm indicado que o uso das regras de concordância se tem incrementado sobretudo na fala dos jovens, em função do deslocamento dos indivíduos para trabalhar nos grandes centros urbanos, da ação dos meios de comunicação de massa e da escola, é possível que o uso da regra de concordância de número com predicativos e particípios passivos venha em breve a fazer parte do repertório da gramática da comunidade, mesmo que com uma frequência de aplicação ainda baixa.

sujeito, ou seja, praticamente nulo. Pode-se então pensar em um paralelo entre as duas estruturas, já que, de acordo com o modelo da Regência e Ligação (CHOMSKY, 1981, 1986), os predicativos também estariam em adjunção ao SN sujeito, que se moveria para posição de sujeito do verbo de ligação, em estrutura superficial, para receber o caso nominativo. As duas estruturas são representadas nas árvores abaixo:

(8) Representações estruturais da adjunção do adjetivo ao nome núcleo do SN e da adjunção do predicativo ao sujeito, na estrutura profunda das orações com verbos de ligação:

Esse paralelo estrutural explica por que o nível de aplicação da regra de concordância nominal é o mesmo nos dois casos. Resta saber se esse paralelo estrutural se confirma no caso da concordância de gênero.

15.3.2 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro

Em relação à concordância de gênero com predicativos e particípios passivos, não é possível fazer-se uma comparação entre a variação que se observa no português afro-brasileiro e outras variedades do português brasileiro, porquanto não são conhecidas análises sobre esse tópico no Brasil. Essa ausência pode ser justificada pela impressão de que não ocorre uma variação significativa na grande maioria das variedades do português brasileiro. Isso reforça a ideia de que a categoria gramatical do gênero foi menos afetada pelo processo de transmissão linguística irregular no Brasil, pelas razões expostas acima.

Nas amostras de fala vernácula das três comunidades rurais afro-brasileiras isoladas do Estado da Bahia, foram depreendidas 460 ocorrências de predicativos e particípios de estruturas passivas ligados a sujeitos femininos, como em (9) a (12):

(9) Naquele tempo, **eu** era **muntcho tímida**. (falante do sexo feminino)

(10) **A casa de farinha** tá **bom**.

(11) **A lavôra** foi **plantada**.

(12) **A associação**... foi **fundado**... eu nem sei mais!

Nessa base de dados, a regra de concordância foi aplicada em 373 ocorrências (cf. exemplos (9) e (11)), com 87 ocorrências de falta de concordância (cf. exemplos

(10) e (12)). Isso perfaz uma frequência de aplicação da regra de concordância de 81% dos casos. Como o nível de aplicação da regra de concordância de gênero no interior do SN na comunidade de Helvécia foi da ordem de 95%, comprova-se, mais uma vez, que o nível de aplicação das regras de concordância nominal é mais elevado no interior do SN do que em estruturas de predicação, o que corrobora a ideia de que há mais coesão estrutural naquelas estruturas do que nestas. Por outro lado, confirma-se também o paralelismo estrutural apresentado em (8) acima, pois a frequência de marcação do feminino em adjetivos, possessivos e modificadores, em geral colocados à direita do núcleo no interior do SN, na gramática da comunidade de fala de Helvécia, é de 76% do total de ocorrências (LUCCHESI, 2000, p.261); ou seja, um nível bastante próximo ao da concordância em estruturas de predicação.

Buscando-se identificar os fatores da estrutura linguística que condicionavam o fenômeno variável em estudo, foram definidas as seguintes variáveis explanatórias: (i) tipo de predicação (predicado nominal, predicado verbo-nominal ou estrutura passiva); (ii) tipo de sujeito (sujeito simples ou composto, sujeito não realizado, sujeito pronominal, sujeito sem o núcleo realizado, ou sujeito atualizado por um pronome relativo); (iii) ordem dos constituintes (sujeito-predicado ou predicado-sujeito); (iv) presença ou ausência de material interveniente entre o verbo e o predicativo ou o particípio passado; (v) aplicação ou não da regra de concordância nominal de gênero no SN sujeito; (vi) aplicação ou não da regra de concordância verbal; (vii) tipo de indicação do gênero no sujeito; (viii) referência ao falante no sujeito; (ix) caracterização semântica do sujeito em termos de sua animacidade; (x) estrutura sintagmática do predicativo (adjetivo, substantivo, possessivo, demonstrativo, etc.); e (xi) tipo de atributo do predicativo (essencial ou acidental). Dessas, o programa de regras variáveis VARBRUL selecionou como estatisticamente relevantes as seguintes: (i) aplicação ou não da regra de concordância nominal de gênero no SN sujeito; (ii) tipo de indicação do gênero no sujeito; e (iii) referência ao falante no sujeito. Das variáveis sociais consideradas (idade, sexo, nível de escolaridade e estada fora da comunidade), apenas a idade foi selecionada. As subseções seguintes contêm os resultados das variáveis explanatórias selecionadas como estatisticamente relevantes.

15.3.2.1 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro relacionada com a concordância nominal de gênero no SN sujeito

A correlação entre a aplicação da regra de concordância nominal nos predicativos e particípios passivos e a aplicação dessa mesma regra no SN sujeito, que especifica o número e o gênero desses predicativos e particípios, está

fundamentada teoricamente no *princípio da coesão estrutural* proposto por Lucchesi (2000, p.141-144).[8] Esse princípio prediz que a probabilidade de o falante aplicar a regra de concordância com os predicativos e particípios passados é maior quando ele aplica a regra de concordância no SN sujeito que especifica as categorias gramaticais de gênero e de número desses predicativos e particípios (cf. exemplos (13) e (14) abaixo), do que quando isso não ocorre (cf. exemplos (15) e (16)). Os resultados do processamento quantitativo, apresentados na Tabela 2 confirmaram essa hipótese.

(13) **Uma perna** dela ficô **dura**.

(14) Eu tenho **uma tia** que é... que foi **ofendido**.

(15) **Esse** é **neta** de Maria Santa.

(16) **As coisa tudo** inda era mais **custoso**.

Tabela 2 - Aplicação da regra de concordância de gênero em predicativos e particípios segundo a aplicação da regra de concordância de gênero no SN sujeito

CONCORDÂNCIA NO SN SUJEITO	Nº DE OC. / TOTAL	FREQ.	P. R.
Com concordância no SN sujeito	154/192	80%	**.52**
Sem concordância no SN sujeito	07/13	54%	**.21**
Total	161/205	79%	—

Nível de significância: .039.

Os resultados constantes na Tabela 2 demonstram que a aplicação da regra de concordância com predicativos e particípios cai significativamente quando essa regra não é aplicada no SN sujeito; passando de 80%, no primeiro caso, para 54%, no segundo (resultados confirmados na proporção dos pesos relativos: .52 *versus* .21). Tais resultados fornecem mais sustentação empírica para o princípio da coesão estrutural no processamento linguístico dos falantes, de acordo com a visão da concorrência de gramáticas (cf. seção 4.2.2 no capítulo 4 deste livro).

15.3.2.2 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro relacionada com a indicação do gênero no SN sujeito

A proposição dessa variável deriva de uma extensão do *princípio da saliência fônica* (NARO; LEMLE, 1976) já consagrado na pesquisa variacionista no Brasil. Desse modo, a probabilidade de aplicação da regra de concordância de gênero com os predicativos e particípios seria maior, quanto mais saliente fosse a indicação do gênero no sujeito com o qual esses predicativos e particípios se relacionam. A

[8] Cf. seção 4.2.2 no capítulo 4 deste livro.

análise baseou-se na seguinte escala decrescente de saliência na indicação do gênero: (i) gênero indicado lexicalmente (*mãe, cabra, Maria,* etc.); (ii) gênero marcado morfologicamente, em nomes ou determinantes e modificadores (*professora, irmã, menina,* etc., ou *ela, essa, nossa,* etc.); e (iii) sem indicação do gênero, lexical ou morfológica, seja no nome núcleo do sujeito, seja nos pronomes e determinantes (*perna, chuva, coisa,* etc., ou *eu, a gente, nós,* etc.). A prevalência da indicação lexical sobre a marcação morfológica apoia-se no fato de que, em processos de reestruturação gramatical decorrentes de contato entre línguas, o primeiro parâmetro prevalece sobre o segundo. Assim, nos crioulos de base portuguesa da África, a morfologia do gênero é totalmente eliminada, sendo substituída pela indicação lexical, por meio de palavras tipo 'fêmea/macho' ou 'homem/mulher', nos casos em que a indicação do gênero é relevante (cf. seção 12.3 no capítulo 12 deste livro). Por outro lado, a distinção entre os nomes e determinantes aplica-se sobretudo para os casos em que o núcleo do SN sujeito não é realizado, como no exemplo (15) acima. Os resultados dessa variável são apresentados na Tabela 3.

Tabela 3 - Aplicação da regra de concordância de gênero em predicativos e particípios segundo a indicação do gênero no SN sujeito

INDICAÇÃO DO GÊNERO	Nº DE OC. / TOTAL	FREQ.	P.R.
mãe, cabra, Maria	41/42	98%	**.90**
professora, irmã, menina	23/25	92%	**.74**
ela, essa, nossa	99/109	91%	**.66**
eu, a gente, nós	60/68	88%	**.37**
perna, chuva, coisa	142/199	71%	**.32**
Total	365/443	82%	—

Nível de significância: .039.

Ao comprovar a escala hierárquica proposta, os resultados proporcionam um significativo paralelo entre a situação linguística das comunidades rurais afro-brasileiras isoladas e as situações de contato entre línguas em que se formaram os crioulos de base lexical portuguesa na África. Em ambos os casos, a indicação lexical predomina sobre a marcação morfológica. A diferença estaria então no grau de erosão da estrutura morfológica, pois, enquanto nos crioulos ocorreu uma eliminação total da morfologia flexional, nas comunidades rurais afro-brasileiras desencadeou-se um processo de variação que se estende por todo o paradigma flexional, não ocorrendo a eliminação de suas unidades, o que possibilitaria a emergência de processos de gramaticalização — fator determinante para a reestruturação original de uma nova gramática: a gramática da língua crioula qualitativamente distinta da gramática de sua língua lexificadora. Essa reestruturação a partir dos processos de gramaticalização constitui, portanto, um divisor de águas entre as situações prototípicas de crioulização e as situações de transmissão linguística irregular mais leves. No caso dos crioulos,

há uma diferença qualitativa em relação à língua-alvo. No segundo caso, a diferença é quantitativa, pois, por mais elevado que seja o nível de variação no uso dos morfemas flexionais, essa variação não atingiu o nível crítico que possibilita a criação de novos elementos gramaticais.

Voltando aos resultados da Tabela 3, observa-se uma inversão entre os nomes e pronomes ou determinantes marcados morfologicamente e os nomes e pronomes ou determinantes não marcados morfologicamente. No plano dos elementos marcados morfologicamente, os nomes favorecem mais a concordância do que os pronomes ou determinantes; entre os elementos não marcados, ocorre o contrário. Isso pode ser explicado, porque os nomes que se flexionam quanto ao gênero possuem o traço semântico [+humano] ou [+animado], enquanto os nomes que não se flexionam quanto ao gênero exibem, em sua grande maioria, o traço semântico [–animado]. Já em relação aos pronomes e determinantes, a situação é diferente, os pronomes que não se flexionam em gênero (e.g., *eu, você, nós, a gente*) possuem o traço semântico [+humano], ao passo que os pronomes e sobretudo os determinantes que se flexionam em gênero (e.g., *ele/a, esse/a, nosso/a*) podem-se referir a seres inanimados. Portanto, pode-se ver aí a influência do traço semântico [+humano] favorecendo a concordância — como já demonstrado na tradição dos estudos variacionistas —, mesmo que o grupo de fatores que focalizava esse parâmetro estrutural não tenha sido selecionado pelo programa das regras variáveis. Além disso, o próximo grupo de fatores selecionado pelo programa relaciona-se especificamente a esse aspecto.

15.3.2.3 A variação na concordância nominal de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro relacionada com a referência ao falante no sujeito da oração

Foi estruturado também um grupo de fatores com base no fato de o sujeito que especifica o gênero dos predicativos e particípios passivos referir-se ou não ao próprio falante, conforme exemplificado a seguir.

- **o sujeito inclui o falante**

(17) **Eu** fiquei junto com eles sossegad**a**.

- **o sujeito não inclui o falante**

(18) **Ela** é assim meia bestinh**a**.

Como se pode observar em (18), o fato de o sujeito não se referir ao próprio falante não significa necessariamente que esse sujeito tenha o traço semântico [–humano]. Não obstante, todas as ocorrências de sujeito [–humano] ou [–animado] incluem-se nesse segundo fator do grupo. Os resultados desse grupo de fatores são apresentados na Tabela 4.

Tabela 4 - Aplicação da regra de concordância de gênero em predicativos e particípios segundo a referência ao falante no sujeito da sentença

REFERÊNCIA AO FALANTE	Nº DE OC. / TOTAL	FREQ.	P. R.
Inclui o falante	58/66	88%	**.70**
Não inclui o falante	315/394	80%	**.46**
Total	373/460	81%	—

Nível de significância: .039.

Os resultados indicam que o falante tende a realizar mais a concordância de gênero quando se refere a si próprio, um parâmetro que não costuma ser considerado nas análises variacionistas, mas que pode revelar um tipo de **condicionamento egocêntrico** no comportamento linguístico. Por outro lado, esse quadro também se ajusta às frequências observadas com a variável traço semântico animacidade, em que os predicativos e particípios passivos que se ligam a sujeitos com o traço semântico [+animado] superam os predicativos e particípios passivos que se ligam a sujeitos com o traço semântico [–animado] em onze pontos percentuais na frequência de aplicação da regra de concordância de gênero. Entretanto, esse grupo de fatores não foi selecionado como estatisticamente relevante pelo programa das regras variáveis.

15.3.2.4 Encaixamento da variação na concordância de gênero em estruturas passivas e de predicativo do sujeito na estrutura social das comunidades rurais afro-brasileiras

A única variável social selecionada pelo programa das regras variáveis como estatisticamente relevante foi a faixa etária dos informantes, cujos resultados são apresentados na Figura 1.

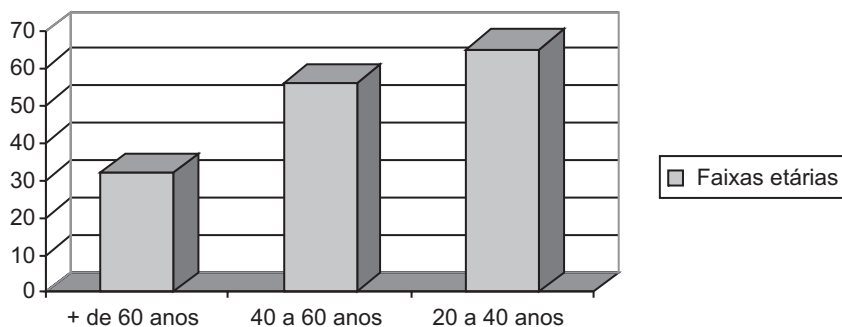



Figura 1 - Nível de aplicação da regra de concordância de gênero em estruturas passivas e de predicativo do sujeito no português afro-brasileiro segundo a faixa etária do falante

Como se pode observar, o nível de uso da regra de concordância de gênero com predicativos e particípios passivos aumenta progressivamente entre os falantes mais jovens, nas comunidades rurais afro-brasileiras. Esse padrão ascendente é um indicador de uma mudança em curso, através da qual o uso da regra de concordância está sendo implementado nos padrões de comportamento linguístico dessas comunidades. Confirma-se, assim, um quadro de nivelamento linguístico, em que os padrões de uso da língua nessas comunidades rurais estaria se alterando em função de influências externas, sobretudo da difusão dos padrões de comportamento linguístico dos grandes centros urbanos. Os vetores desse nivelamento linguístico seriam: a influência dos meios de comunicação de massa; o deslocamento de indivíduos que buscam trabalho nas grandes cidades; e a ação da escola, em que pese a precariedade do sistema de educação pública.

Apesar de não terem sido selecionadas pelo programa de cálculo probabilístico, vale a referência aos percentuais das demais variáveis sociais, que corroboram, em grande parte, essa interpretação. A frequência de uso da regra de concordância entre os indivíduos que viveram fora da comunidade por pelo menos seis meses supera a frequência de uso daqueles que não se ausentaram da comunidade em cinco pontos percentuais: 83% *versus* 78%. A grande maioria desses indivíduos que deixam a comunidade vão trabalhar nos grandes centros urbanos, colocando-se na liderança de um processo de nivelamento linguístico da cidade para o campo. Os indivíduos que têm qualquer contato com a escolarização tendem a usar mais a regra de concordância do que os que são completamente analfabetos, na mesma proporção: 84% *versus* 79%. Isso poderia ser o indicador de que o contato com a escola modificaria o uso linguístico em relação às regras de concordância.

Por fim, o resultado da variável sexo não correspondeu à expectativa construída nas análises anteriores sobre as regras de concordância nas comunidades rurais afro-brasileiras. Essas análises têm constatado que os homens tendem a ter um comportamento mais próximo do padrão culto do que as mulheres, diferentemente do que se tem observado nas análises sociolinguísticas feitas nos grandes centros urbanos de países plenamente industrializados, que identificam as mulheres como um segmento mais sensível às formas de prestígio (CHAMBERS; TRUDGILL, 1980, p.97-98; LABOV, 1981, p.184). Essa inversão se explicaria, no universo sociocultural das comunidades rurais do interior do Brasil, pelo fato de que os homens têm normalmente mais contato com o mundo exterior, enquanto as mulheres se mantêm mais restritas ao universo rural e doméstico (LUCCHESI, 2000, p.288-290). Entretanto, os resultados desta análise sobre a concordância de gênero com predicativos e particípios passivos revelou um padrão de quase equilíbrio entre homens e mulheres, sendo que essas usariam as regras numa frequência ligeiramente superior: 82% *versus* 80%, dos homens.

15.4 Conclusão

Esta análise variacionista sobre a concordância nominal em estruturas sentenciais da voz passiva e de predicativo do sujeito demonstrou, em primeiro lugar, que o nível de variação no uso das regras de concordância nessas estruturas é bem mais elevado nas comunidades rurais afro-brasileiras do que na fala dos segmentos de escolaridade média dos grandes centros urbanos brasileiros. No que concerne à regra de concordância de número, as análises de Scherre (1991) e Dias (1996) apontam para uma frequência de uso da regra da ordem de 50%, enquanto esta regra está praticamente ausente da gramática das comunidades de fala rurais afro-brasileiras. Já em relação à regra de concordância de gênero, enquanto a própria inexistência de estudos sobre esse tópico para as variedades linguísticas urbanas brasileiras indica a falta de um nível significativo de variação, foi constatado um nível de variação da ordem de 20% nas comunidades rurais afro-brasileiras.

Partindo-se do princípio de que essa diferença se explica pelo fato histórico de que as comunidades rurais afro-brasileiras teriam sido mais diretamente afetadas pelo contato entre línguas do que as variedades urbanas do português brasileiro, pode-se concluir que essa interferência do contato entre línguas afetou mais a categoria gramatical do número que a categoria gramatical do gênero. Isso poderia ser explicado em função do caráter mais gramatical e redundante da primeira na estrutura linguística do português.

No plano dos condicionamentos linguísticos, traçou-se um significativo paralelo estrutural, com base na coincidência da frequência entre a concordância dos modificadores em adjunção à direita do nome núcleo no interior do SN e a concordância com os predicativos e particípios passivos. Esse paralelo é válido, tanto para a categoria gramatical do número, quanto para a do gênero. No tocante ao encaixamento linguístico da variação, verificou-se, em primeiro lugar, que o fenômeno é condicionado pela aplicação da regra de concordância de gênero no SN sujeito que especifica o gênero do predicativo e do particípio passivo; o que se explica pelo princípio da **coesão estrutural**, com base no fato de que o falante tende a fazer mais a concordância com o predicativo ou o particípio quando realiza a concordância no SN sujeito do que quando não a realiza.

O consagrado princípio da **saliência fônica** também se mostrou relevante, pois o falante faz mais a concordância com os predicativos e particípios, quanto mais saliente for a indicação do gênero no SN sujeito. Nessa escala de saliência, observou-se uma supremacia da indicação lexical sobre a marcação morfológica, o que permite traçar um paralelo entre o processo de estruturação gramatical das comunidades de fala afro-brasileiras e o processo bem mais radical de reestruturação gramatical dos crioulos de base lexical portuguesa da África, em que a indicação

lexical do gênero substitui a marcação morfológica, eliminada na situação de contato radical em que essas línguas se formaram. Por fim, constatou-se também que o falante tende a fazer mais a concordância quando fala de si próprio, o que parece revelar um **condicionamento egocêntrico** no comportamento linguístico dos falantes, até então pouco explorado pela pesquisa sociolinguística.

Num plano mais geral, pode-se dizer que este estudo da variação no uso das regras de concordância em estruturas de predicativos e particípios passivos integra um amplo círculo de estudos sobre as regras de concordância que têm contribuído para construir um significativo panorama sociolinguístico acerca desse aspecto da gramática do português brasileiro. Scherre e Naro (1997, p.107), ao apresentar os resultados das variáveis sociais de suas análises sobre a concordância verbal, a concordância de número no interior do SN e a concordância de número com predicativos e particípios, identificaram um quadro de variação estável no universo de falantes urbanos com nível de escolarização médio. Já as análises desses e outros tópicos relativos à concordância nominal e verbal em comunidades rurais afro-brasileiras isoladas têm demonstrado inequivocamente uma tendência ao incremento do uso das regras de concordância nessas comunidades (cf. capítulos 11, 12 e 14 deste livro). O contraste desses resultados ajusta-se perfeitamente à visão da polarização da realidade linguística brasileira desenvolvida por Lucchesi (1994, 1998, 2001a, 2002a, 2002b, 2006a). Essa polarização atualiza-se na tendência de variação estável identificada nos padrões de comportamento linguístico dos falantes da norma semiculta da amostra Censo do PEUL, no Rio de Janeiro, *vis-à-vis* ao processo de mudança em direção ao incremento da regra de concordância em estruturas passivas e de predicativo do sujeito nas comunidades rurais afro-brasileiras isoladas que a análise variacionista apresentada neste capítulo revelou.

16 O emprego do modo subjuntivo

Vivian Meira

A variação no uso dos modos, especificamente do subjuntivo, em português é um fato incontestável. Nos padrões de fala das comunidades rurais afro-brasileiras isoladas do interior do país, também se observa essa variação no uso dos modos verbais. Entretanto, consideramos que, diferentemente do que ocorre em outras variedades do português, a variação no uso do subjuntivo nessas comunidades está diretamente relacionada com o processo de transmissão linguística irregular, desencadeado pelo massivo contato do português com as línguas africanas, ocorrido nos períodos colonial e imperial. Nessas comunidades, diferentemente do que se observa em pesquisas no português urbano, o subjuntivo vem gradativamente ganhando ambientes antes ocupados apenas pelo indicativo, visto que os antepassados desses falantes, que aprenderam o português precariamente na condição de escravos, devem ter adquirido inicialmente as formas do indicativo, pois este modo, por se referir a eventos reais, tende a ser mais usado na comunicação, podendo ser definido, portanto, como o modo morfologicamente não marcado. Nesse sentido, cremos que o subjuntivo vem sendo adquirido por estes falantes em decorrência da difusão dos meios de comunicação e de toda a infraestrutura propiciada pela urbanização do país.

A partir dos pressupostos da teoria da variação linguística (LABOV, 1972a, 1994), estudamos o uso do modo subjuntivo em quatro comunidades rurais afro-brasileiras do interior do Estado da Bahia: as comunidades de Cinzento, Helvécia, Barra e Bananal (Rio de Contas) e Sapé. Foram analisadas 28 entrevistas, com informantes distribuídos por sexo (masculino e feminino), idade (faixa 1: de 20 a 40 anos; faixa 2: de 41 a 60 anos; faixa 3: de 61 a 80 anos; e faixa 4: com mais de 80 anos), escolaridade (analfabeto e semianalfabeto) e comunidade (considerando a variação em cada comunidade). Para cumprir os seus objetivos, este capítulo está estruturado da seguinte maneira: a primeira seção apresenta uma revisão do tema em foco, o modo subjuntivo; a segunda seção apresenta os resultados da análise variacionista do uso do modo subjuntivo considerando o tipo de oração em que a forma verbal é empregada: as orações relativas, as orações completivas e as orações adverbiais; a terceira seção apresenta uma análise de conjunto do emprego

do modo subjuntivo, reunindo todas as ocorrências da base de dados e abordando o fenômeno variável tanto em seu encaixamento linguístico, quanto no encaixamento social nas comunidades rurais afro-brasileiras analisadas; na última seção, estão as nossas considerações finais.

16.1 O modo subjuntivo

Etimologicamente, o termo *subjuntivo* significa 'subordinado', do latim *subjunctivus,* e, literalmente, serve para ligar, para subordinar, por isso tem seu emprego, por excelência, numa oração subordinada. De acordo com Cunha e Cintra (1985, p.466), o nome subjuntivo denota que "uma ação, ainda não realizada, é concebida como dependente de outra, expressa ou subentendida. Daí o seu emprego normal na oração subordinada". No entanto, esse modo verbal não é empregado em todos os casos de subordinação; ao contrário, em muitas orações subordinadas usa-se preferencialmente o indicativo, e o modo subjuntivo é empregado, muitas vezes, em orações denominadas independentes ou absolutas.

A tradição gramatical apresenta um sistema de modo verbal, cujo emprego se baseia, ora em critérios semânticos, ora em critérios sintáticos e formais. Além disso, os modos verbais são também condicionados por regras facultativas cuja aplicação é regulada por fatores intencionais e subjetivos, pois, muitas vezes, cabe à atitude do falante o emprego de determinado modo, mesmo que a estrutura gramatical indique o uso de um modo específico. Por outro lado, podemos verificar também que a complexidade em estudar a forma verbal se dá inicialmente a partir do fato de que um mesmo morfema acumula em si as noções de tempo e de modo. Semanticamente, são valores distintos, mas são indissociados morfologicamente.

De acordo com o critério semântico, a tradição gramatical estabelece uma relação inequívoca entre a atitude subjetiva do falante frente ao fato e o modo verbal empregado, firmando uma oposição semântica entre o indicativo e o subjuntivo, de forma que é inerente a este uma atitude de incerteza e àquele uma atitude de convicção acerca do real. Mas nem sempre essa relação é de fato inequívoca. Com base em Santos (2003), buscamos observar o que existe de comum em: *Quero que você **estude** hoje*, *Duvido que você **estude** hoje*, ***Estude** hoje* ou *Fico contente que você **estude** hoje*. Curiosamente, poderíamos nos perguntar o que permite a mesma forma — ***estude*** — assumir conteúdos semânticos distintos, como dúvida, vontade, sentimento, etc. Levando-se em conta o que prescreve a tradição gramatical no que diz respeito ao emprego do subjuntivo, diríamos que tal modo é determinado automaticamente pelo tipo de verbo da oração principal e, assim, deparamo-nos com um "problema" sintático. Por outro lado, assume-se também que o modo

verbal é dependente de uma atitude do falante diante de um fato ou de uma proposição enunciada, referindo-se, assim, a uma questão semântico-pragmática. No entanto, o emprego do subjuntivo não se deve exclusivamente a uma questão sintática ou semântico-pragmática, mas também, e com certa frequência, a expressões de dúvida, a conjunções, a advérbios, etc. Daí poderíamos supor que os princípios sintático e semântico-pragmático de emprego do subjuntivo estariam estreitamente relacionados com tais *partículas*.

Portanto, podemos dizer que uma explicação meramente sintática não abarcaria todas as ocorrências do subjuntivo, motivo pelo qual recorre-se a critérios puramente semânticos, que, por sua vez, são insuficientes, não cobrindo todos os aspectos de uso desse modo em português, valendo-se, assim, de definições de *varia* ordem.[1]

16.2 Análise variacionista do uso do modo subjuntivo no português afro-brasileiro, considerando-se o tipo de oração em que a forma verbal é empregada

Em nossa análise, o uso do modo subjuntivo é dividido em função das distintas estruturas sintáticas em que ele pode ser empregado, estabelecendo-se três grandes contextos: (i) o uso do modo subjuntivo em orações relativas, (ii) o uso do modo subjuntivo em orações completivas e (iii) o uso do modo subjuntivo em orações adverbiais ou junto a expressões de dúvida. Assumindo-se que o uso do modo subjuntivo é condicionado de forma distinta em cada um desses contextos no plano da estrutura linguística, a análise variacionista de cada um desses contextos foi feita separadamente, e os seus resultados serão apresentados nas seções seguintes deste capítulo.

---

[1] Nesse sentido, estamos de acordo com Santos (2003, p.34) que afirma: "Surgem então vários paradoxos difíceis de aceitar. Fala-se num modo subordinado e, muitas vezes, sem sentido próprio, reflexo automático de uma regra gramatical destituída, até, de força real, mas, ao mesmo tempo, enumeram-se todos os sentidos possíveis das expressões regentes, considerando que o [subjuntivo] surge por força desses mesmos sentidos. Fala-se em ocorrência obrigatória, decorrente de uma regência pré-estabelecida, mas não se põe de parte que o [subjuntivo] possa exprimir um sentido próprio. Fala-se nos casos em que um falante *escolhe* voluntariamente o [subjuntivo] e, ao fazê-lo, cria uma diferença pertinente em relação a outras formas que poderiam surgir no mesmo contexto, mas também se insiste no seu caráter excepcional, que não invalidaria a regra geral".

16.2.1 Análise variacionista do modo subjuntivo nas orações relativas

As variáveis linguísticas explanatórias utilizadas para a análise do *uso do subjuntivo em orações relativas*[2] foram as seguintes: (i) localização temporal do evento expresso na oração relativa em relação ao momento da enunciação; (ii) tempo do subjuntivo previsto no uso culto e (iii) morfologia verbal.

Com a variável *localização temporal do evento expresso em relação ao momento da enunciação,* levamos em consideração a hipótese de que uma referencialidade posterior ao momento da enunciação, por se relacionar com eventos irreais e hipotéticos e, portanto, com o valor semântico do subjuntivo, tenda a favorecer o uso desse modo verbal.

Com as variáveis *tempo do subjuntivo previsto no uso culto e morfologia verbal,* procuramos avaliar a atuação do princípio da saliência fônica, isto é, se as formas mais marcadas foneticamente na oposição subjuntivo *versus* indicativo favorecem o uso do subjuntivo. Tal princípio foi observado tanto na diferença entre as formas do presente e do imperfeito do subjuntivo, quanto com relação à questão da regularidade e irregularidade dos verbos.

Por outro lado, como se trata de comunidades que apresentam um passado marcado pelo contato entre línguas, em que se operam processos de simplificação morfológica (cf. capítulo 3 deste livro), esperamos que as marcas do tempo futuro dos verbos regulares sejam mais recorrentes, visto que estas se assemelham às formas do infinitivo, o que teria facilitado a sua aquisição. Nesse sentido, observe-se que a marca do futuro tende a se assemelhar à do infinitivo, mesmo em verbos irregulares, como em:

(1) É aonde nós tamos por aí até o dia que Deus **querê**. (RC-24)[3]

16.2.1.1 Localização temporal do evento expresso na oração relativa em relação ao momento da enunciação

A partir dos resultados da Tabela 1, podemos observar que o uso do subjuntivo é largamente favorecido quando o evento referido na relativa se localiza em um momento posterior ao momento da ilocução, confirmando assim a nossa hipótese (cf. 16.2.1).

---

[2] Em função das limitações de espaço, serão apresentados apenas os resultados das variáveis explanatórias mais relevantes. Uma análise detalhada do fenômeno pode ser encontrada em Meira (2006).

[3] Nos exemplos, à indicação das comunidades (CZ = Cinzento; HV = Helvécia; RC = Rio de Contas, SP = Sapé), segue-se o número do informante, como registrado nos *corpora* base. A metodologia de constituição dessas amostras de fala é descrita no capítulo 5 deste livro.

Tabela 1 - Uso do subjuntivo nas orações relativas no português afro-brasileiro segundo a localização temporal do evento expresso na oração relativa

LOCALIZAÇÃO TEMPORAL	Nº DE OC. / TOTAL	FREQ.	P. R.
Posterior à ilocução	17/31	61%	**.93**
Anterior à ilocução	12/66	15%	**.37**
Simultaneamente à ilocução	09/38	13%	**.36**
Total	**38/135**	**28%**	—

Nível de significância: .044.

Tanto os pesos relativos quanto os valores percentuais revelam que o uso do subjuntivo é desfavorecido quando os eventos referidos são anteriores ou simultâneos ao momento da ilocução, contextos exemplificados em (2) e (3), respectivamente.

(2) As comida *que num* ***dava*** pá comê, num podia come. (CZ-09)

(3) É difici ir assim alguém *que num* ***usa*** o chapéu. (CZ-01)

Isso se ajusta ao valor semântico do subjuntivo, que se relaciona com eventos hipotéticos e irreais, mais frequentes na referência a eventos posteriores ao momento da enunciação, como em:

(4) ... ali as coisa que *a pessoa pensá em comprá*, tá tudo ali, né. (RC-05)

De fato, esse plano do *irrealis* está mais diretamente ligado ao futuro do que ao presente e ao passado; na verdade, os eventos que se situam no futuro são objetivamente irreais, por maior que seja a certeza do falante de sua realização.

16.2.1.2 Tempo do subjuntivo previsto no uso culto

Tomando por base o uso culto, foi estudado, no português afro-brasileiro, o uso dos tempos presente, imperfeito e futuro do subjuntivo, exemplificados em (5), (6) e (7), respectivamente:

(5) Uma quantidadizinha que **dê** pra... num compra mais maior quantidade que o dinheiro não dá. (SP-05)

(6) ... ela mesmo fazia a recepção pra que as pessoa que **quisesse** tomá dinhêro... (CZ-11)

(7) ... tudo que a gente **tivé**, assim, a gente leva... (HV-01)

Os resultados dessa variável estão apresentados na Tabela 2.

Tabela 2 - Uso do subjuntivo nas orações relativas no português afro-brasileiro segundo o tempo do subjuntivo previsto no uso culto

FORMA PREVISTA NO USO CULTO	Nº DE OC. / TOTAL	FREQ.	P. R.
Futuro do Subjuntivo	17/31	55%	**.78**
Imperfeito do Subjuntivo	09/38	24%	**.46**
Presente do Subjuntivo	12/66	18%	**.38**
Total	**38/135**	**28%**	—

Nível de significância: .044.

Observamos que o grande favorecedor do uso das formas do subjuntivo é o futuro, enquanto o imperfeito fica um pouco abaixo da média geral de uso (24% contra 28%, do geral), desfavorecendo ligeiramente o emprego do subjuntivo (p.r. de .46). O contexto de presente é aquele que mais desfavorece o uso desse modo verbal, com apenas 18% de frequência de uso e peso relativo de .38.

Segundo Mattoso Câmara Jr. (1979), não existia o tempo futuro do subjuntivo no latim vulgar e esse tempo verbal se originou de formas flexionais volitivas e subjuntivas; a partir daí, desenvolveu-se um futuro modal, que conduziu as línguas românicas a um futuro temporal. Lyons (1979) afirma que o caráter modal do tempo futuro, existente desde o latim, atua no futuro temporal das línguas românicas, o que pode levar a uma associação entre os contextos de emprego deste e do subjuntivo. A ideia de projeção futura desencadeada pelo tempo futuro relaciona-se com o traço *irrealis*, na medida em que o futuro indica apenas uma suposição, hipótese ou, como afirma Mattoso Câmara Jr. (1970 [2002]), o tempo futuro, assim como o pretérito, mantém uma oposição em orações que designam uma condição prévia do que será dito, pois um evento futuro sugere que poderá acontecer ou não. Além disso, as formas do futuro em sua grande maioria coincidem com as formas do infinitivo, o que facilitaria a sua aquisição. Já o imperfeito e o presente, que apresentam morfemas exclusivos, seriam mais lentamente incorporados ao uso da comunidade de fala. E, entre esses dois, as formas do imperfeito levariam vantagem por apresentarem um morfema foneticamente mais saliente e regular, o *-sse-* (que possui o padrão silábico CV). Por outro lado, a alternância da vogal temática que marca as formas do presente do subjuntivo seria a de mais difícil aquisição.

16.2.1.3 Morfologia verbal

Na análise da variável *morfologia verbal*, foi feita a distinção entre verbos regulares e irregulares, como exemplificado em (8) e (9), respectivamente:

(8) Mas a criação que 'ocê... 'ocê **sustentá** na mão, 'cê é obrigado tê a mandioca. (CZ-12)

(9) Se eu topá ôta pessoa que me **dá** assistênça e me ajuda é o pai a mesma coisa. (CZ-06)

A expectativa é que o subjuntivo fosse mais usado com os verbos irregulares, em função da maior quantidade de material fonético envolvido na flexão modo-temporal, numa extensão da aplicação do princípio da saliência fônica. No entanto, os verbos regulares favorecem mais o uso das formas do subjuntivo do que os verbos irregulares, como podemos verificar na Tabela 3.

Tabela 3 - O uso do subjuntivo nas orações relativas no português afro-brasileiro segundo a morfologia verbal

FLEXÃO VERBAL	Nº DE OC. / TOTAL	FREQ.	P. R.
Regular	14/50	28%	**.66**
Irregular	24/112	21%	**.42**
Total	**38/162**	**23%**	—

Nível de significância: .044.

Os resultados apontam no sentido de que a regularidade favorece mais a aquisição das formas do subjuntivo do que a saliência fônica, que favoreceria o uso das formas irregulares. Desse modo, o princípio da saliência fônica não teve sua aplicação confirmada pelos dados, não contribuindo para a aquisição das marcas de subjuntivo.

16.2.2 Análise variacionista do modo subjuntivo nas orações completivas

Com o descarte das ocorrências em que o uso do indicativo em orações completivas era categórico, como, por exemplo, em orações selecionadas por verbos declarativos (cf. “Elas **disse que de caminhão elas num anda** nunca” (CZ-01)) e perceptivos (cf. “Isso é pra **vê se eu aprendo** ainda alguma coisa” (CZ-01), a base de dados restante ficou constituída por apenas 80 ocorrências em que o uso do subjuntivo era variável, não possibilitando a obtenção de resultados consistentes no nível da análise probabilística do Programa das Regras Variáveis (VARBRUL). Por isso, os resultados apresentados serão baseados apenas na frequência, expressa nos resultados percentuais.

Tendo isso em vista, do total de 80 ocorrências de orações completivas dessa base de dados, foram registradas apenas 23 ocorrências de formas do subjuntivo, o que corresponde a praticamente 30% do total. Esse uso do modo subjuntivo foi condicionado pelas variáveis: (i) tipo da oração em que a completiva está encaixada; (ii) avaliação do falante acerca do nível de realidade do evento referido na oração completiva; (iii) tempo do subjuntivo previsto no uso culto e (iv) morfologia verbal.

Com as variáveis *tipo da oração em que a completiva está encaixada* e *avaliação do falante acerca do nível de realidade do evento referido na oração completiva*, procuramos verificar a relação entre o modo subjuntivo e o valor semântico de irrealidade

contido na oração principal, ou seja, se o contexto semântico do evento referido na oração principal tende a favorecer a aquisição das formas de subjuntivo pelos falantes das comunidades de fala analisadas. O esperado é que esse modo verbal tenda a ocorrer em completivas encaixadas em orações que, em linhas gerais, contenham proposições hipotéticas, que estejam sob o domínio da dúvida e da incerteza e, portanto, que estejam associadas ao valor semântico do subjuntivo.

16.2.2.1 Tipo da oração em que a completiva está encaixada

No que concerne à variável *tipo da oração em que a completiva está encaixada*, amalgamamos os fatores, reunindo-os em apenas três: afirmativo, negativo e a junção entre os fatores condicional e oração com verbo modal, exemplificados em (10), (11) e (12), respectivamente:

(10) Espero que daqui pra frente, a gente... a situação da gente **vai melhorá** mais... cada vez mais. (CZ-03)

(11) Eles num gostam que **ande entrano** no mato pa caçá não? (SP-04)

(12) Você também foi errado, se você num quisesse que ela **fosse**, você ficava queto, num mandava ela arrumá. (SP-01)

Levando-se em conta o fato de o modo subjuntivo estar associado a hipóteses e eventos duvidosos e irreais, a expectativa era a de que os falantes tendessem a fazer uso do subjuntivo quando a oração principal apresentasse uma condição hipotética sobre um evento; o que também se aplica às negativas, que modificam a modalidade da predicação, contrariando-a ou contradizendo-a (MATEUS et al., 2003). Por outro lado, as orações afirmativas parecem estar mais relacionadas com eventos reais, o que poderia contribuir para desfavorecer o uso das formas do subjuntivo. Os resultados são apresentados na Tabela 4.

Tabela 4 - Uso do subjuntivo nas orações completivas no português afro-brasileiro segundo o tipo de oração em que a completiva está encaixada

TIPO DE ORAÇÃO	Nº DE OC. / TOTAL	FREQ.
Condicional	03/07	43%
Negativa	06/18	33%
Afirmativa	14/55	25%
Total	23/80	29%

Como podemos constatar, quando está dependente de uma oração condicional ou de uma oração negativa (com 43% e 33% de uso do subjuntivo, respectivamente), a completiva favorece mais o uso do subjuntivo do que quando está subordinada a uma oração afirmativa, em que o uso do subjuntivo ocorre com uma frequência de apenas vinte e cinco pontos percentuais. Esses resultados confirmam a relação do modo subjuntivo com o valor semântico de irrealidade, pois, estando essas orações

completivas encaixadas em orações negativas e condicionais, elas tendem a forçosamente conter proposições hipotéticas ou contrafactuais.

16.2.2.2 Tempo do subjuntivo previsto no uso culto

A análise da variável *tempo do subjuntivo previsto no uso culto*, nas orações completivas, contou apenas com ocorrências do presente e do imperfeito (exemplificados abaixo), pois não foram encontradas ocorrências do futuro nesse contexto na amostra de fala analisada:

(13) Tá difícil... tá difícil... tá... e essas aí, eles num qué que **tire** não. (HV-07)

(14) Eu queria que **estudasse**, eu tinha dois menino, os dois menó tá estudano. (SP-05)

A expectativa era a de que a forma mais perceptível do imperfeito favorecesse mais o uso das formas do subjuntivo, como já visto acima. Os resultados dessa variável são apresentados na Tabela 5.

Tabela 5 - Uso do subjuntivo nas orações completivas no português afro-brasileiro segundo o tempo do subjuntivo previsto no uso culto

CONTEXTO DE USO	Nº DE OC. / TOTAL	FREQ.
Contexto de Imperfeito do Subjuntivo	11/33	33%
Contexto de Presente do Subjuntivo	11/45	24%
Total	**22/78**	**28%**

Confirmamos a aplicação do princípio da saliência fônica, visto que a alta frequência de uso do subjuntivo nos contextos de imperfeito se deve à maior força morfofonológica desse tempo verbal, pois o morfema do imperfeito *-sse-* apresenta um padrão CV mais consistente em termos de seu material fonético do que a alternância vocálica que indica o presente do subjuntivo.

16.2.2.3 Morfologia verbal

Como já visto acima, a variável *morfologia verbal* diz respeito à diferença de material fônico nas formas dos verbos regulares e irregulares no uso do modo subjuntivo. A Tabela 6 apresenta os resultados dessa variável.

Tabela 6 - Uso do subjuntivo no português afro-brasileiro segundo a morfologia verbal da oração completiva

TIPO MORFOLÓGICO DO VERBO	Nº DE OC. / TOTAL	FREQ.
Irregular	15/49	31%
Regular	08/30	27%
Total	**23/79**	**29%**

A partir dos resultados, verificamos que, diferentemente do que ocorreu nas relativas, o princípio da saliência fônica prevaleceu sobre a regularidade no processo de aquisição das formas do subjuntivo nas comunidades rurais afro-brasileiras isoladas.

16.2.3 Análise variacionista do modo subjuntivo nas orações adverbiais

Quanto ao uso do subjuntivo em orações adverbiais, analisamos as seguintes variáveis linguísticas: (i) nível de realidade do evento contido na oração adverbial; (ii) localização temporal do evento expresso na oração adverbial em relação ao momento da enunciação; e (iii) morfologia verbal.

16.2.3.1 Nível de realidade do evento contido na oração adverbial

Essa foi a segunda variável selecionada pelo programa; esperávamos verificar quais os contextos semânticos poderiam em maior intensidade influenciar o uso do subjuntivo. Foram postulados cinco situações: fato irreal, hipotético, pressuposto, contrafactual e ocorrido, respectivamente, exemplificados abaixo:

(15) Eu acho que, se eu **tivesse estudado** mais, era importante pra mim... (CZ-03).

(16) Se **visse** qualqué coisa de errado, ele trazia pra secretaria, aí tomava "xepo". (SP-01).

(17) Eu quero construí um casa cá em cima. Quano eu **fô fazê** cá, que eu vô fazê, mas, por enquanto, na que a gente mora num tem. (SP-04).

(18) Quando chega aí o período de julho, agosto, setembro... só fica as galha sêca, aí **se 'cê num tivé** o capim pro bicho comê, aí ele... 'cê acaba perdeno as criação. (CZ-03).

(19) Levantava aqui cedinho aqui, fazia café **pra que eu fosse sete**, seis e meia eu já tava lá na casa deles, pra num perdê a hora. Aí trabaiava o dia todo (CZ-01).

A nossa expectativa era a de que as formas de subjuntivo fossem mais recorrentes nos contextos de irrealidade. A Tabela 7 apresenta os resultados obtidos na quantificação dos dados.

Como podemos ver a partir dos resultados, o contexto de *irrealidade*, apresentando 93% de frequência de subjuntivo, valor este muito acima da média geral (53%), com um peso relativo de .90, pode ser considerado um fator favorável ao uso desse modo verbal. Em seguida, o nível *hipotético*, com um peso relativo de .51 e uma frequência de 66%, também demonstra condicionar o uso das formas de subjuntivo nas comunidades de fala analisadas. Por outro lado, o nível *ocorrido* é

aquele que mais desfavorece o uso do subjuntivo, com peso relativo de .25 e uma frequência de 16%, muito abaixo da média geral, o que confirma a associação entre o uso das marcas do subjuntivo e o valor semântico de irrealidade, visto que eventos ocorridos supõem em princípio que já foram realizados e tendem a se associar ao valor [+*realis*], inerente ao indicativo. Já os níveis *pressuposto* e *contrafactual* revelaram-se fatores neutros em relação ao uso do subjuntivo, com um peso relativo de, respectivamente, .49 e .47. De certa forma, eventos pressupostos denotam o entendimento do falante e, portanto, o que este considera como verdade sobre determinado fato. Confirma-se, assim, o condicionamento do uso das formas de subjuntivo pelo parâmetro semântico, visto que as formas de subjuntivo, nos contextos marcados pelo traço de irrealidade, vêm ganhando espaço em detrimento das formas do modo indicativo.

Tabela 7 - Uso das formas do modo subjuntivo em orações subordinadas adverbiais e junto a expressões de dúvida no português afro-brasileiro segundo a variável nível de realidade do evento contido na oração adverbial

NÍVEL DE REALIDADE	Nº DE OC. / TOTAL	FREQ.	P. R.
Irreal	14/15	93%	**.90**
Hipotético	282/426	66%	**.51**
Pressuposto	107/306	35%	**.49**
Contrafactual	67/121	55%	**.47**
Ocorrido	4/25	16%	**.25**
Total	**474/893**	**53%**	—

Nível de significância: .017.

16.2.3.2 Localização temporal do evento expresso na oração adverbial em relação ao momento da enunciação

Como observado em relação às orações relativas, adotou-se a hipótese de que a referência a um evento posterior ao momento da enunciação tende a favorecer o uso do subjuntivo, o que se confirmou na análise quantitativa dos dados, como se pode ver na Tabela 8.

Tabela 8 - Uso das formas do modo subjuntivo em orações subordinadas adverbiais e junto a expressões de dúvida no português afro-brasileiro segundo a variável localização temporal do evento expresso na oração adverbial em relação ao momento da enunciação

LOCALIZAÇÃO TEMPORAL	Nº DE OC. / TOTAL	FREQ.	P. R.
Posterior	344/548	63%	.67
Anterior	100/205	49%	.47
Simultâneo	30/141	21%	.07
Total	474/894	53%	—

Nível de significância: .017.

Os eventos anteriores ou simultâneos ao momento da enunciação desfavorecem o uso do subjuntivo nas comunidades de fala analisadas, apresentando, respectivamente, pesos relativos de .47 e .07. Por outro lado, quando o evento referido na adverbial é posterior, o uso das formas subjuntivas é maior, com peso relativo de .67, o que se relaciona com o valor semântico do subjuntivo. Esses resultados estão de acordo com aqueles encontrados com as orações relativas, ratificando a ideia de que o uso do subjuntivo é também condicionado por um parâmetro semântico.

16.2.3.3 Morfologia verbal

Como já visto, essa variável testa a hipótese de que uma diferença fonética inerente às formas do subjuntivo dos verbos irregulares tende a favorecer o uso das formas desse modo verbal. Na Tabela 9, podemos verificar os resultados dessa variável.

Tabela 9 - Uso das formas do modo subjuntivo em orações subordinadas adverbiais e junto a expressões de dúvida no português afro-brasileiro segundo a variável morfologia verbal

FLEXÃO VERBAL	Nº DE OC. / TOTAL	FREQ.	P. R.
Regular	218/353	62%	.60
Irregular	256/541	47%	.43
Total	474/894	53%	—

Nível de significância: .017.

Assim como nos resultados encontrados com as orações relativas, os verbos regulares tendem a favorecer a aquisição das formas subjuntivas, prevalecendo a regularidade sobre o princípio da saliência fônica na aquisição das formas do subjuntivo nas comunidades rurais afro-brasileiras. Nesse sentido, deve-se levar em consideração o fato de que, em situações de contato, é comum a eliminação da alomorfia, que caracteriza a flexão dos verbos irregulares, o que, de certa forma, faz com que os falantes não flexionem os verbos irregulares, fazendo uso algumas vezes do infinitivo desses verbos, em vez da forma flexionada (e.g., *se a gente* ***ir*** *hoje, amanhã a gente tamo de volta.* (HV-04); *se ele* ***querê*** *vim fazê pouso longe...* (HV-20)).

16.3 Análise de conjunto da variação no emprego do modo subjuntivo no português afro-brasileiro

Encontramos uma disparidade entre os resultados registrados por Pimpão (1999), ao estudar o subjuntivo no português urbano, e os resultados que

obtivemos, diferença essa que dimensiona a distância que separa o português urbano culto e o português rural afro-brasileiro. A diferença na frequência de uso do subjuntivo entre essas duas variedades do português brasileiro pode passar dos sessenta pontos percentuais, como se pode ver na Tabela 10.

Tabela 10 - Frequência geral de uso das formas do modo subjuntivo no tempo presente em orações subordinadas adverbiais, completivas e relativas no português urbano e no português afro-brasileiro

FORMAS DO SUBJUNTIVO	PORTUGUÊS URBANO	PORTUGUÊS AFRO-BRASILEIRO
Orações completivas	84%	24%
Orações relativas	82%	18%
Orações adverbiais	79%	52%

Essa significativa diferença entre os percentuais constitui mais uma evidência empírica da polarização sociolinguística no português do Brasil, como sistematizada por Lucchesi (2001, 2002a, 2002b, 2006). E uma observação mais detalhada no quadro da variação no uso do subjuntivo pode revelar, para além da diferença quantitativa, uma diferença qualitativa nas tendências de mudança que se delineiam em cada uma dessas duas variedades linguísticas brasileiras.

No português afro-brasileiro, as formas do subjuntivo parecem estar penetrando em contextos em que, no passado, só ocorriam as formas do indicativo, configurando-se uma tendência de aquisição das formas do subjuntivo na gramática das comunidades rurais afro-brasileiras isoladas. Esse processo de mudança seria o resultado da crescente influência linguística dos grandes centros urbanos sobre todas as regiões do interior do país, que se intensifica a partir da segunda metade do século XX. Com isso altera-se a realidade pretérita dessas comunidades. Formadas em situações de contato maciço entre o português e as línguas africanas, essas comunidades passaram por um processo de transmissão linguística irregular, que se caracteriza principalmente pela simplificação morfológica (LUCCHESI, 2003b e capítulo 3 deste livro). Assim, essas comunidades, em sua formação, devem ter reduzido drasticamente o uso do subjuntivo, generalizando-se o uso das formas morfologicamente não marcadas do indicativo.

16.3.1 O encaixamento linguístico

Com o objetivo de cotejar os resultados quantitativos do uso do subjuntivo nas diversas estruturas sintáticas em que esse modo verbal pode ser empregado, estruturou-se uma variável com os seguintes fatores: oração subordinada adverbial condicional, oração subordinada adverbial temporal, demais tipos de oração subordinada adverbial, oração optativa, oração com *talvez* e demais advérbios e

expressões de dúvida, oração subordinada completiva e oração relativa. Os resultados desse grupo de fatores são apresentados na Tabela 11.

Tabela 11 - Uso das formas do modo subjuntivo em orações subordinadas completivas, relativas e adverbiais junto a expressões de dúvida no português afro-brasileiro segundo a variável tipo de oração em que está inserido o verbo

TIPO DE ORAÇÃO	Nº DE OC. / TOTAL	FREQ.	P. R.
Optativa	25/32	78%	**.81**
Subordinada adverbial condicional	352/513	69%	**.71**
Subordinada adverbial	13/29	45%	**.47**
Subordinada adverbial temporal	82/300	27%	**.31**
Orações completivas	22/78	28%	**.31**
Orações relativas	37/157	24%	**.26**
Com *talvez* e expressão de dúvida	02/19	11%	**.10**
Total	**533/1128**	**47%**	—

Nível de significância: .016.

Os resultados demonstram que as formas de subjuntivo são mais usadas em orações optativas, com uma frequência de 78% do total de ocorrências, contra 47% de frequência geral, e um peso relativo de .81, bem acima dos .50, que indica neutralidade do fator. Isso se deve a motivações semânticas, visto que as orações optativas exprimem desejos do locutor e, dessa forma, algo ainda não realizado, ou seja, refere-se ao valor [–*realis*] inerente ao subjuntivo.

A relação do uso do subjuntivo em proposições com o traço semântico [–*realis*] foi ratificada também quando observamos o uso desse modo verbal em orações subordinadas condicionais, com uma frequência de 69%, peso relativo de .71, já que estas, em linhas gerais, expressam uma hipótese, valor semântico associado ao subjuntivo.

Por outro lado, as orações com expressões de dúvida são as que mais desfavorecem o uso das formas de subjuntivo. Encontramos 11% do total de ocorrências, muito abaixo da média geral de 47%, com peso relativo de .10, bem abaixo também dos .50. Nesse caso, acreditamos que se observa uma lógica funcional de eliminação da redundância, pois o advérbio *talvez* e demais expressões de dúvida denotam o valor semântico [–*realis*], que seria expresso igualmente pela forma verbal no subjuntivo. Já nas orações optativas, em que foi registrado o maior uso das marcas de subjuntivo, não há advérbios ou expressões que marquem dúvida ou incerteza, o que leva essa marca a ser usada no verbo. A eliminação de formas redundantes reduzindo a morfologia flexional do verbo é comum nas línguas crioulas, e a história das comunidades analisadas é caracterizada pelo contato entre línguas e pelo processo de transmissão linguística irregular, que teria levado à eliminação de marcas redundantes.

As orações completivas (p. r. .31) e relativas (p. r. .26) também tendem a desfavorecer o uso das formas subjuntivas quando comparado com o uso dessas formas nas orações adverbiais.

16.3.2 As variáveis sociais

A análise do encaixamento social do uso do subjuntivo nas comunidades rurais afro-brasileiras isoladas foi feita tomando-se como base empírica o conjunto total de ocorrências das três variáveis linguísticas definidas para a análise do encaixamento linguístico, ou seja, reunindo todos os contextos em que as formas do modo subjuntivo foram empregadas em uma única rodada do VARBRUL. Nessa rodada, as variáveis sociais *faixa etária*, *nível de escolaridade* e *comunidade de origem* foram selecionadas como estatisticamente relevantes. Nas seções seguintes, serão discutidos os resultados quantitativos de cada uma dessas variáveis.

16.3.2.1 Faixa etária

Com a variável *faixa etária*, procuramos observar se os falantes mais jovens são responsáveis pela aquisição das formas de subjuntivo, apresentando maior percentual de uso. Vejamos os resultados obtidos com essa variável.

Tabela 12 - Uso das formas do modo subjuntivo em orações subordinadas completivas, relativas e adverbiais junto a expressões de dúvida no português afro-brasileiro segundo a variável faixa etária

FAIXA ETÁRIA	Nº DE OC. / TOTAL	FREQ.	P. R.
20 a 40 anos	131/263	50%	**.55**
41 a 60 anos	188/350	54%	**.56**
61 a 80 anos	144/341	42%	**.45**
81 anos em diante	71/175	41%	**.40**
Total	**534/1129**	**47%**	—

Nível de significância: .016.

Com base nos dados das quatro comunidades, verificamos um processo de aquisição das marcas de subjuntivo, uma vez que os falantes das faixas 1 e 2 apresentam maior peso relativo de uso desse modo verbal com, respectivamente, pesos relativos de .55 e .56, e os falantes mais velhos, das faixas 3 e 4, são responsáveis pelo menor peso relativo, com .45 e .40, respectivamente, o que demonstra que há um processo de incremento das formas do subjuntivo nas comunidades de fala analisadas.

De certo modo, a mudança deve ter tido início por volta de 1965, em decorrência das alterações sociais ocorridas nas comunidades, como o contato com diferentes grupos sociais, com os meios de comunicação de massa, com a

escola, etc., e afetou principalmente as faixas 1 e 2, os mais jovens, que tendem atualmente a uma maior realização desse modo verbal. Por outro lado, os falantes das faixas 3 e 4 usam com pouca frequência o subjuntivo, o que estaria relacionado com um estado anterior da língua, influenciado pelo contato entre línguas e pelo processo de transmissão linguística irregular.

Rocha (1997) apresenta, em seu estudo sobre o subjuntivo no português urbano, algumas questões no sentido de que a alternância entre subjuntivo e indicativo pode constituir um fenômeno de variação em que o subjuntivo está perdendo ambiente para o indicativo ou de uma variação que envolve a alternância entre os dois modos, verificando, assim, a partir de seus resultados, que está ocorrendo uma alternância entre os modos. No entanto, temos de afirmar que isso parece não funcionar no português afro-brasileiro, uma vez que neste está havendo o inverso: o indicativo está perdendo (aos poucos) ambiente para o subjuntivo, pois este modo vem sendo gradativamente adquirido pelos membros dessas comunidades, conforme demonstrado pelos resultados da variável *faixa etária*.

A partir dos resultados obtidos nas comunidades de fala analisadas, as formas do subjuntivo estão sendo aos poucos assimiladas pelos membros dessas comunidades, situação inversa à do meio urbano, em que as formas do subjuntivo estão cedendo lugar às formas do indicativo. Essas distintas tendências de mudança, que se observam em duas variedades do português brasileiro, constituem um dos parâmetros cruciais que Lucchesi (1998, 2001, 2002a, 2002b e 2006a) considerou para a caracterização da polarização sociolinguística da realidade linguística brasileira, dividida em duas grandes normas: a culta e a popular, que, para além das diferenças nas frequências de uso das variáveis linguísticas, apresentariam também tendências distintas de mudança linguística e sistemas igualmente diferenciados de avaliação social das variantes linguísticas. A questão da avaliação social das variantes é diretamente afetada pela atuação da escola, como veremos a seguir.

16.3.2.2 Nível de escolaridade

Apesar de não constituir um estereótipo no Brasil, o uso do subjuntivo parece ser afetado pela ação da escola. Pimpão (1999), em sua pesquisa sobre o uso do presente do modo subjuntivo na cidade de Florianópolis, distribuiu os falantes, quanto à escolaridade, em primário, ginásio e colegial e observou que os falantes dos níveis escolares mais elevados usam mais as formas do subjuntivo. Nesta análise, os informantes das quatro comunidades foram divididos em dois grupos: o dos analfabetos e o grupo dos semianalfabetos, aqueles que tinham qualquer contato com a escolarização, até mesmo aqueles que sabiam apenas assinar o nome.

Os resultados indicam que os semianalfabetos usam mais o subjuntivo (p. r. .60) do que os analfabetos (p. r. .43), demonstrando que a escolarização do meio rural interfere nos padrões de uso desse modo verbal.

Tabela 14 - Uso das formas do modo subjuntivo em orações subordinadas adverbiais e junto a expressões de dúvida no português afro-brasileiro segundo a variável nível de escolaridade

NÍVEL DE ESCOLARIDADE	Nº DE OC. / TOTAL	FREQ.	P. R.
Semianalfabeto	239/482	50%	.60
Analfabeto	295/647	46%	.43
Total	534/1128	47%	—

Nível de significância: .016.

16.3.2.3 Comunidade de origem

Como a hipótese de nosso trabalho é a de que o contato entre línguas afetou o processo de desenvolvimento da língua portuguesa no Brasil, então, os efeitos mais salientes desse contato seriam mais notáveis em comunidades cuja origem estaria ligada a agrupamentos de ex-escravos e/ou de escravos foragidos e que até recentemente têm se mantido em um grau relativo de isolamento. Sendo assim, em comunidades rurais afro-brasileiras isoladas, podemos identificar reflexos de fenômenos que seriam condicionados pelo contato entre línguas e que podem contribuir para definir as características atuais da língua falada no Brasil, especialmente em sua variedade rural.

Cremos que o isolamento da comunidade tende a conservar alguns fenômenos variáveis na língua. Considerando-se a formação histórica e social das comunidades rurais de Cinzento, Helvécia, Rio de Contas e Sapé, observamos certa uniformidade, visto que todas estão relacionadas com situações de contato entre línguas e com a fuga da escravidão, mantendo-se ainda em certo isolamento em relação ao meio urbano. No entanto, os contextos atuais dessas comunidades não são idênticos, pois há aquela que apresenta vestígios da existência de um falar crioulo, a comunidade de Helvécia, e a que está mais sujeita à influência externa, a de Rio de Contas, em função do turismo (cf. capítulo 2 deste livro). Nesse sentido, esperamos que as formas de subjuntivo fossem menos usadas na comunidade de Helvécia. Observemos a Tabela 15.

Tabela 15 - Uso das formas do modo subjuntivo em orações subordinadas completivas, relativas e adverbiais junto a expressões de dúvida no português afro-brasileiro segundo a variável comunidade de origem

COMUNIDADES	Nº DE OC. / TOTAL	FREQ.	P. R.
Sapé	120/214	56%	**.60**
Rio de Contas	78/168	46%	**.51**
Cinzento	180/393	46%	**.50**
Helvécia	155/352	44%	**.44**
Total	**534/1129**	**47%**	—

Nível de significância: .016.

A partir dos resultados, podemos observar que em Sapé foi registrado um maior uso das formas de subjuntivo em relação às outras comunidades, com um peso relativo de .60, o que se ajusta à localização dessa comunidade, situada no Recôncavo Baiano, próximo à cidade do Salvador. Na verdade, a proximidade geográfica da comunidade com um grande centro urbano está causando modificações na estrutura linguística, influenciando variações e/ou mudanças no sentido da norma urbana culta. Em Rio de Contas, registramos um peso relativo de .51, o que pode ser explicado pelo contato com padrões linguísticos propiciados pelo turismo no município. Cinzento está praticamente na média geral de uso do subjuntivo (p. r. .50). Com relação a Helvécia, observamos que esta é a comunidade que apresenta o menor índice de uso do subjuntivo (p. r. .44), uma vez que a história dessa comunidade está relacionada com um passado crioulo (cf. capítulo 2 deste livro).

16.4 Considerações finais

Para compreendermos o português do Brasil, é necessário conhecer a história tanto do português urbano, quanto do português rural, observando a origem e a constituição dessas realidades linguísticas. Os africanos e, sobretudo, os seus descendentes, foram um importante agente na difusão do português, na formação da sociedade brasileira (MATTOS E SILVA, 2004). Os escravos africanos adquiriram o português de forma irregular, sem normatização, produzindo uma variedade da língua portuguesa marcada pela redução na morfologia flexional. Na medida em que esse modelo defectivo foi se convertendo na língua materna dos seus descendentes, foram se formando comunidades de fala bastante distantes do português brasileiro dito culto no que concerne ao uso de morfemas flexionais. Esse cenário do português rural tem se alterado desde meados do século XX, em decorrência do êxodo rural e da ação dos meios de comunicação de massa, verificando-se um processo de nivelamento linguístico a partir dos padrões que emanam das grandes cidades.

Um reflexo desse processo está no quadro de variação no uso do modo subjuntivo nas comunidades rurais afro-brasileiras isoladas, cujo padrão ascendente observado no cotejo das gerações indica um incremento das formas do subjuntivo na gramática dessas comunidades de fala. Assim, diferentemente do que ocorre nas variedades urbanas do português brasileiro, a morfologia do subjuntivo, que havia se reduzido drasticamente em função do processo de transmissão linguística irregular que marca a formação dessas comunidades, estaria sendo reintroduzida em seus padrões coletivos de uso da língua, em função de influências externas, sobretudo dos grandes centros urbanos, através dos meios de comunicação de massa.

Esse processo de mudança, no sentido de uma ampliação do uso do modo subjuntivo no plano da estrutura linguística, seria determinado por dois fatores. O primeiro fator é de natureza morfológica, em que a forma mais saliente, em termos morfofonológicos (tanto os verbos quanto os tempos), favorece a implementação das formas do subjuntivo, podendo-se estabelecer dois estágios: um primeiro, em que os falantes tendem a fazer uso das formas de subjuntivo mais regulares; um segundo, em que os falantes, ao adquirir alguma consciência sobre as formas padrão, atentam mais para as formas de subjuntivo em que o material fônico é mais perceptível. O segundo fator é semântico: as formas do subjuntivo começam a ser empregadas com referência a eventos claramente irreais, já que o subjuntivo é associado ao traço semântico *irrealis*. Acreditamos que o princípio da transparência semântica pode explicar o incremento das formas do subjuntivo, a partir do momento em que o falante percebe uma oposição entre um modo relacionado com o *realis* e outro associado ao *irrealis*, passando a dispor de diferentes meios expressivos para efetivar a comunicação.

17 O objeto direto anafórico: a categoria vazia e o pronome lexical

*Cristina Figueiredo**

O português falado no Brasil (PB) conta com duas estratégias não padrão de realização do objeto direto anafórico de terceira pessoa (ODA): a categoria vazia e o pronome lexical *ele*. Trabalhos realizados sobre o tema em *corpora* sincrônicos e diacrônicos atestam que essas variantes são condicionadas, principalmente, por fatores semânticos, e que a implementação da categoria vazia (CV) se deu a partir do século XVIII, em decorrência do uso da CV com antecedente sentencial, em razão da perda progressiva do clítico neutro de terceira pessoa, estendendo-se, mais tarde, a outros contextos. A perda dos clíticos referenciais de terceira pessoa (*o/a*) leva à ampliação do uso da CV, e se reflete no aparecimento do pronome lexical *ele* (PL) como objeto direto (TARALLO, 1993; DUARTE, 1986; CYRINO, 1997; FIGUEIREDO, 2004, dentre outros).[1]

Os exemplos abaixo ilustram as duas possibilidades:

(1) a. Maria comprou o disco de Roberto Carlos, mas Ana não quis ouvir $____i$. **(CV)**

b. João está fazendo muita falta aqui. Vou procurar **ele** em todo lugar. **(PL)**

Há ainda uma outra forma de realização do ODA, a retomada através de um item lexical (SN), que pode ocorrer através da repetição do mesmo item lexical, como em (2a) ou por outro item que remeta ao antecedente, como em (2b).

(2) a. Há duas semanas, baixou na embaixada de Portugal, em Brasília, **um rapaz**$_i$ pedindo asilo político. [...] Finalmente, horas depois, o embaixador convenceu **o rapaz**$_i$ a pedir ajuda na sede da CNBB, vizinha da embaixada portuguesa. (*Veja*, 21 set. 2005, p.20.)

b. Na festa, o aniversariante ganhou **o novo livro de João Ubaldo Ribeiro**$_i$ e não vê a hora de começar a ler **a obra**$_i$.

---

* Agradeço a Ilza Ribeiro pela leitura atenta e pelas discussões realizadas durante a elaboração deste texto.

[1] Não discuto aqui o estatuto das realizações de *ele* como objeto no português arcaico. Veja-se Martínez (1993), para uma análise detalhada desta questão.

Este capítulo apresenta um estudo dos fatos apontados, as possibilidades de realização do ODA, a partir de dados do *Corpus* Base do Português Afro-Brasileiro do Estado da Bahia do Projeto *Vertentes*, contemplando as 24 entrevistas de quatro comunidades rurais afro-brasileiras do interior do Estado da Bahia: Helvécia, localizada no Extremo Sul; Rio de Contas, na Chapada Diamantina; Cinzento, no Semiárido; e Sapé, no Recôncavo Baiano. A escolha de tal universo de observação se baseia na hipótese de que o contexto de multilinguismo existente no Brasil, principalmente durante os três primeiros séculos de colonização, teria sido o fator que desencadeou, no PB, características sintáticas que o diferenciam do português europeu (PE) (cf. capítulo 1 deste livro). Sabe-se que, no PE, só as formas clíticas e a CV, com algumas restrições,[2] são possíveis como realização de ODA, como pode ser observado na distribuição a seguir:

(3) a. Muitos dentifrícios protegem contra as bactérias, mas está clinicamente provado que só Colgate Total protege contra as bactérias abaixo da gengiva. Pensamos que lhe dá proteção mais completa. Exprimente-**o**. (CYRINO 2001, p.177-178[3]) (**CL**)[4]

b. O pirata partiu para as Caraíbas depois de ter guardado $____{i.}$ cuidadosamente no cofre. - sobre o tesouro (RAPOSO, 2004, p.44) (**CV**)

(√PB/?PE[5])

c. Ganhei muitas laranjas. Vou descascar elas para fazer um doce (**PL**)

Para verificar a interferência do multilinguismo na realização das estratégias de retomada do ODA no português afro-brasileiro, este capítulo está estruturado da seguinte maneira: a primeira seção apresenta uma revisão dos estudos desse fenômeno em dialetos urbanos e os principais fatores de condicionamento das variantes; na seção 17.2, são analisados os dados quantitativos do dialeto rural afro-brasileiro, apontando-se os fatores linguísticos e extralinguísticos relevantes no uso da CV e do PL; a seção 17.3 contém os resultados da análise dos condicionamentos sociais; e, na seção 17.4, são apresentadas as conclusões a que se chegou a respeito da variante eleita pelos falantes durante o período de contato linguístico, bem como a trajetória do PL no dialeto analisado.

---

[2] Cf. Raposo 1986, 2004, e referências citadas.

[3] Exemplo do português europeu.

[4] Clítico acusativo de 3ª pessoa.

[5] Raposo (1986) considera que a CV em contexto de ilha, como em 3b, seja agramatical, porém Raposo (2004) reconsidera o juízo de gramaticalidade anteriormente proposto.

17.1 Explicitando as variantes: uma revisão

Esta seção apresenta sucintamente resultados de pesquisas sobre o estado do ODA em *corpora* de amostras dos dialetos paulista e carioca, precisamente os utilizados por Duarte (1986) e por Averburg (2000); e estabelece uma avaliação das variantes de acordo com a tradição gramatical. Tem como objetivo principal delinear um quadro geral das variantes (padrão e não padrão) em uso no Brasil, o que fornecerá elementos para entender as questões de transmissão linguística irregular (cf. capítulo 3 deste livro).

A variante CL, embora considerada padrão, é pouco encontrada no vernáculo brasileiro. A pesquisa variacionista empreendida por Duarte (1986) sobre as estratégias de realização do uso do ODA num dialeto urbano de São Paulo comprova que o uso do clítico de terceira pessoa (*o/a*) está condicionado, principalmente, pela escolaridade, e que essa variante é preferida em situações mais formais e nos textos escritos. A importância do fator escolaridade pode ser observada na Figura 1, de Averburg (2000), que estuda a realização variável do objeto direto anafórico em textos escritos de estudantes do Rio de Janeiro.

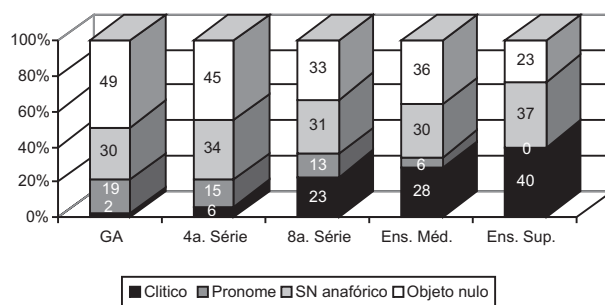



Figura 1 - O objeto direto anafórico segundo a escolaridade no dialeto carioca
Fonte: Averburg, 2000, p.58.

Nota-se, a partir dos resultados numéricos da Figura 1, que quanto maior a escolaridade, maior é o percentual de uso do CL, 40% das ocorrências em textos de universitários, enquanto a variante mais estigmatizada, o PL, não é encontrada nesse grupo.

Outra estratégia bastante comum de manter a coesão entre os enunciados de um texto, função fundamental dos ODA, é a retomada de antecedentes através do uso de um SN. Essa estratégia, que Fávero (1993, p.23) chama de coesão referencial por reiteração, pode ser realizada através de vários recursos lexicais, tais como nos exemplos apresentados pela autora, transcritos a seguir:

(4) a. Repetição do SN:
O fogo acabou com tudo. **A casa** estava destruída. D**a casa** não sobrara nada.

b. Sinonímia:

A criança caiu e chorou. Também, **o menino** não ficava quieto.

c. Expresssões referenciais definidas:

O cantor Sting tem lutado pela preservação da Amazônia. **O ex-líder da banda Police** chegou ao Brasil ontem.

A frequência de realização da variante SN não é bastante significativa nos *corpora* estudados por Duarte (1986), não alcançando 20% dos dados, mas deve ser considerada na investigação do ODA, uma vez que, para os dialetos em que não há a forma padrão, CL, essa é uma forma neutra ao padrão linguístico estabelecido pela tradição gramatical. Segundo Duarte (1986), a frequência dos SNs aumenta em grupos com maior escolaridade e faixa etária mais alta, superando o uso da variante PL, em função do estigma que ainda recai sobre esta última. O gráfico de Averburg (2000), do dialeto carioca (Figura 1), mostra uma maior realização dessa estratégia, no patamar de 30% a 40% das ocorrências de ODA, não havendo diferenças substanciais no que diz respeito à escolarização.

Embora considerado não padrão pela tradição gramatical, o pronome *ele,* na posição de objeto direto, tem-se caracterizado como uma marca do PB (cf. exemplo em (1b)). O estudo sincrônico de Duarte (1986) aponta que, inversamente ao comportamento apresentado pelo CL, essa variante é mais frequente entre os falantes mais jovens, e a sua percentagem decresce à medida que a idade do falante aumenta, estando condicionada também pela escolarização. Os números da Figura 1 confirmam os resultados de Duarte (1986), haja vista a sua ausência no grupo com escolaridade máxima, bem como os19% de realização no grupo de alfabetização. Há uma frequência contínua decrescente do PL, de informantes de menor escolarização para aqueles com maior escolaridade, um padrão inverso ao comportamento da variante CL.

A CV[6] é uma estratégia não reconhecida pela norma padrão, nem seu uso sofre qualquer estigma, sendo um recurso presente inclusive em textos escritos em que predomina a norma culta, conforme os exemplos a seguir retirados de revistas:

(5) a. O prefeito de São Paulo José Serra apoiou a candidatura do senador Tasso Jereissati à presidência do PSDB com uma **condição**$_i$. Requisitou a secretaria-geral do partido, o segundo cargo mais importante na hierarquia, para o deputado Eduardo Paes, do Rio de Janeiro. Tasso aceitou ___$_i$. (*Veja*, 12 out. 2005, p.4.)

b. Em Buenos Aires, **Borges**$_i$ está por todo lado. É ler ___$_i$ para crer" (___ = a obra de Borges - o assunto do texto) (*Entre Livros*, jun. 2006, p.39.)

[6] Neste trabalho, considera-se o termo CV de forma ampla, representando qualquer lacuna na posição de OD que esteja em covariação estrita com as demais estratégias. Para a caracterização sintática desta CV, cf. Cyrino (1997), dentre outros.

A Figura 1 revela a importância do fator escolaridade, ao apontar para o declínio no uso da CV: informantes menos escolarizados apresentam 49% de ocorrência, e os mais escolarizados, 23%, uma diferença de quase 30%. Segundo os resultados de Duarte (1986), essa é a variante mais utilizada na retomada do objeto direto no dialeto urbano, e é favorecida, principalmente, pelo traço semântico [–animado] do antecedente. Assim, além do fator escolaridade, outro traço que tem se mostrado determinante na escolha da variante do ODA é a animacidade do seu antecedente (cf. exemplos em (8)).

Ao realizar o inventário das ocorrências para a análise das comunidades de fala do dialeto rural afro-brasileiro, foram consideradas apenas três das estratégias de realização do ODA, a CV, o PL e o SN, visto que a estratégia CL, segundo Figueiredo (2004), não faz parte da gramática dos falantes das comunidades analisadas, uma vez que o uso dessa variante está condicionado ao nível mais alto de escolarização, e essas comunidades são compostas, predominantemente, por analfabetos e semianalfabetos.

17.2 Análise linguística

A análise dos dados segue os pressupostos do modelo variacionista (LABOV, 1972a, 1994), com o emprego do pacote de programas VARBRUL para o seu processamento quantitativo. Foram depreendidas 1.755 ocorrências do ODA, distribuídas pelas três variantes presentes nas comunidades analisadas, segundo a Tabela 1.

Tabela 1 - Distribuição das variantes no dialeto rural afro-brasileiro

VARIANTES	SN	PL	CV	TOTAL
Nº de ocorr.	275	213	1267	1755
%	16	12	72	100

A CV é a estratégia de uso mais difundida na realização do ODA (72%), seguida do SN, com 16%, embora com diferenças quantitativas e percentuais mínimas em relação ao PL (12%). A análise desse quadro geral se mostra mais interessante quando comparado com os resultados de Duarte (1986), do dialeto urbano paulista, representados na Tabela 2.

Tabela 2 - Distribuição das variantes no dialeto urbano paulista

VARIANTES	CL	SN	PL	CV	TOTAL
Nº de ocorr.	79	338	304	1235	1974
%	4.9	17.1	15.4	62.6	

Fonte: Duarte, 1986, p.17.

Nos dois dialetos, rural e urbano, a CV é a estratégia preferida pelos falantes. Diferentemente do dialeto urbano, o CL não é encontrado no dialeto rural, demonstrando que essa é uma estratégia que não faz parte da gramática dos falantes das comunidades aqui estudadas, como já comentado acima; contudo, ressalte-se que, no dialeto urbano, o uso do CL tem baixa frequência (4,9%). A análise conjunta dessas duas estratégias aponta para uma conclusão interessante: a maior realização da CV no dialeto rural (72% contra 62% no dialeto urbano) pode ser um recurso para compensar a perda do CL.

Não há diferenças quantitativas relevantes na estratégia do SN, com 16% das realizações no dialeto rural e 17% no urbano paulista. Contudo, o PL apresenta um percentual menor no dialeto rural do que no urbano, respectivamente 12% e 15%.

Esperava-se que, na fala dos membros das comunidades rurais analisadas, a variante PL fosse mais frequente, já que se partiu do pressuposto de que seu dialeto, tendo em vista o contexto multilíngue em que se formou, apresentasse traços "crioulizantes", como simplificação de formas mais marcadas — o PL [+marcado] para caso nominativo da função sujeito no PE passa a [–marcado] para caso, podendo ser realizado também na função de objeto e substituiria o CL. Contudo, o que se percebe é o desaparecimento da variante [+marcada] para caso acusativo, a forma CL, e o uso frequente da variante neutra, a CV.

Em resumo, os dados computados não permitem dizer que tenha havido, no processo de formação do dialeto rural afro-brasileiro, a substituição da forma CL pelo PL, desde que seu uso é pouco frequente, mas aponta para a substituição do CL pela variante menos marcada, a CV, tendo em vista a sua maior frequência.

Serão apresentados a seguir os fatores relevantes na realização do objeto direto anafórico, segundo a ordem de seleção do programa utilizado para análise quantitativa.

17.2.1 O condicionamento semântico

Este fator tem sido bastante relevante na caracterização da CV em posição de objeto, e muitos autores têm confirmado a preferência pela CV com antecedente [–animado] (DUARTE, 1986; CYRINO, 1997; GALVES, 1984). Exemplos das duas possibilidades estão (6).

(6) a. [–animado]

Eu vem de manhã, ***coloco*** **o fêjão**$_i$ no fogo e deixa ___$_i$ aí. (RC-01)

b. [+animado]

Eles$_i$ (os vereadores) vêm aqui?

Vêm... a gente só ***aborrece*** **eles**$_i$ quando... a vez, se doece um, a gente pede um... um socorro e eles dão, né, (CZ-03)

Os resultados desta pesquisa também revelam que é este o fator mais relevante na escolha das estratégias em foco: a CV é favorecida pelo traço [–animado], com peso relativo de .68, enquanto o traço [+animado] favorece o uso do PL, com peso relativo de .80. Embora este seja um condicionamento bastante significativo, a CV predomina, mesmo nos contextos de traço [+animado], como se pode ver percentualmente na Figura 2.

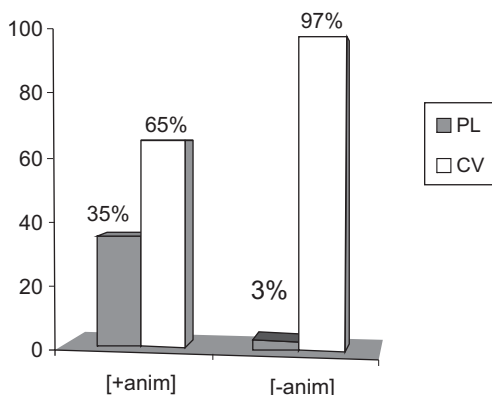



Figura 2 - O traço semântico animacidade do antecedente do ODA no dialeto rural afro-brasileiro

Na seção 17.2.3 este fator será retomado, buscando-se verificar a sua relevância em relação à estrutura do VP.

17.2.2 O paralelismo discursivo

O paralelismo, como definido por Scherre e Naro (1993), reflete a hipótese de que *marcas levam a marcas e zeros levam a zeros*.[7] Segundo os autores, tanto no nível da sentença (paralelismo formal) quanto no nível do discurso (paralelismo discursivo), há correlação entre as marcas explícitas, de modo que *formas gramaticais particulares tendem a ocorrer juntas*. Os exemplos a seguir ilustram a modalidade do paralelismo discursivo, com as duas variantes do ODA, a CV, em (7a), e o PL, em (7b).

(7) a. DOC: E a **cobra**$_i$ pode mordê?

INF: Uai, eles fala que num pode matá ___$_i$, né? Que a gente só passa no qu'é deles.

DOC: Ham, ham.

INF: Se incontrá ____$_i$, cum'é que vai matá ___$_i$?

---

[7] Esta hipótese se baseia em Poplack 1980 apud Scherre e Naro, 1993.

b. DOC: E Jovelina vai também?

INF: Não, **ela** só foi quando a menina casô... quando casô, depois **ela** num foi mais não.

DOC: Num tem vontade de ir lá.

INF: Não, **ela** quase num vai, que sempre quando eu vô dêxo **ela** mais minha mãe, né?

Na análise desse fator, incluiu-se a variante SN, a fim de verificar se o fator paralelismo discursivo atua também em relação à repetição do SN. Os resultados do processamento quantitativo são apresentados na Tabela 3.

Tabela 3 - A estrutura formal do antecedente mais próximo no dialeto rural afro-brasileiro

FORMA DO ANTECEDENTE MAIS PRÓXIMO	SN Nº ocorr. / %	PL Nº ocorr. / %	CV Nº ocorr. / %	TOTAL
SN	102 / **27**	27 / 7%	252 / 66	381
PL	5 / 3	90 / **53**	76 / 44	171
CV	63 / 9	68 / 10	537 / **80**	668
Total	170 / **14**	185 / **15**	865 / **71**	1220

O uso de uma dada variante favorece a sua repetição, em relação à CV (80% nesse contexto *versus* 71% de frequência geral) e ao PL (53% *versus* 15%), ou mesmo o SN (27% *versus* 14%).

17.2.3 A estrutura da frase

A estrutura da frase foi o terceiro fator selecionado pelo VARBRUL como significativo na escolha das variantes, sendo um fator apontado por Duarte (1986) como determinante na escolha da forma ODA no dialeto paulista. As estruturas testadas estão exemplificadas abaixo:

(8) V + OD(SN):

INF: De primêro, eles vinha, mas depois que o pai morreu vêi mais não. Eu sempre vô lá, [**encontro eles**]. (CZ-06)

(9) V + OD(SN) + SP(OI/LOC):

DOC: E me diga uma coisa: E aí vocês fazem como? Cada um paga um poquinho pa comprá a **taça**$_i$, não?

INF: Não, isso aí já vem, é o prefeito mesmo que... que [**manda___**$_i$ **pa gente**].(SP-04)

(10) V + OD(SN) + PRED:

DOC: Graças a Deus. Onde a senhora pega **a lenha**$_i$ pra...?

INF: Não, não, panha assim hoje e dá pa queimá a semana.

DOC: Ah é?

INF: E quando a vez eu acho uma pessoa, eu pago pa [**tira** $____i$ **verde**]. (RC-24)

(11) V + $_{OD}[_{S}SU_{acc}$ + $V_{[-tempo]}]$:

DOC: Cê nunca teve curiosidade de sabê comé que surgiu o Cinzento não?

INF: Sobe não, né?... a vez, eu [**vejo eles contá**], mas num prendei na mente assim não. (SP-03)

No dialeto rural afro-brasileiro, este fator também se mostra relevante, como se pode observar nos números da Tabela 4.

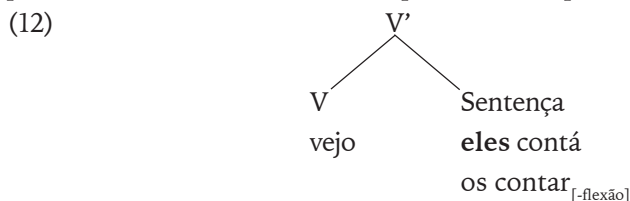
Tabela 4 - A forma do objeto direto anafórico segundo a variável estrutura da frase no dialeto rural afro-brasileiro

ESTRUTURA DA FRASE	PRONOME LEXICAL			CATEGORIA VAZIA		
	Nº ocorr.	%	p.r.	Nº ocorr.	%	p.r.
V+OD(SN)+SP(OI/LOC)[1]	28/292	10	.38	264/292	80	**.62**
V+OD(SN)	152/1120	14	.52	968/1120	86	.48
V+OD(SN)+PRED	11/32	34	**.66**	21/32	66	.34
V+$_{OD}[_{S}SU_{acc}$+$V_{[-tempo]}]$	22/36	61	**.74**	14/36	39	.26
Total	213/1480	14	—	1267/1480	86	—

Na estrutura mais simples, exemplifcada em (8), o peso relativo das variantes CV e PL não são decisivos na escolha dos falantes, pois estão muito próximos ao ponto neutro de .50 (respectivamente, .48 e .52).

O contexto em que o ODA é seguido por um SP(OI/LOC), exemplificado em (9), favorece a CV (p.r. de .62). Já o PL é favorecido, em estruturas em que o ODA está envolvido numa predicação, como nos casos exemplificados em (10) e (11). Em (10), o ODA está inserido numa predicação nominal, formando uma minioração com o predicativo do objeto. No caso da estrutura (11), o ODA, embora receba caso acusativo do verbo da oração principal, expressa o agente do verbo da oração encaixada, mesmo que esse verbo não lhe possa atribuir caso nominativo, pois não está flexionado, como se pode ver na representação feita em (12) abaixo.

(12)

```
            V'
          /    \
         V      Sentença
        vejo    eles contá
                os contar[-flexão]
```

[8] Na posição de SP, foram considerados os termos que exercem a função de objeto indireto e de locativo na posição de argumento do verbo, sendo descartados os termos locativos em posição de adjunção.

Considerando o favorecimento da variante PL nessas estruturas, foi feito o cruzamento dessa variável com a de animacidade. O resultado desse cruzamento pode ser conferido na Tabela 5.

Tabela 5 - Animacidade *versus* estrutura do SV no dialeto rural afro-brasileiro

ESTRUTURA DO SV	[+ANIMADO]		[–ANIMADO]	
	CV	PL	CV	PL
	Nº ocorr / %	Nº ocorr / %	Nº ocorr / %	Nº ocorr / %
V+OD(SN)	246/**67**	120/33	715/96	30/4
V+OD(SN)+SP(OI/LOC)	74/**75**	25/25	188/**98**	3/2
V+OD(SN) + PRED	7/39	11/**61**	14/**100**	—
$V+_{OD}[_{S}SU_{acc}+V_{[-tempo]}]$	13/37	22/**63**	1/**100**	—
Total	340/**66**	178/**34**	918/**97**	33/**3**

Com o cruzamento dos fatores, verifica-se que, quando o traço do antecedente é [+animado], a CV é favorecida em estruturas do tipo V + OD(SN) e V + OD(SN) + SP(OI/ LOC), com 67% e 75% respectivamente. Quando o traço do antecedente é [–animado], é favorecida praticamente por todas as estruturas.

Quanto ao PL, além do traço [+animado] do antecedente, percebe-se que a estrutura complexa é também bastante relevante em sua escolha. Essa variante é favorecida, principalmente, em contextos como aqueles em que o ODA é sujeito da oração encaixada, com 63%, um resultado esperado ao se considerar a hierarquia temática proposta por Baker (1977)[9]. Nessa mesma estrutura, com antecedente [–animado], nenhuma ocorrência com PL foi encontrada, embora exista um dado de CV. Eis o único dado:

(13) INF.: ...aí só é jogá o fogo e jogô a massa, liga o $motô_i$ e [dêxa $____i$ rolá] enté ti... a... a... a farinha tivé pronta]. (SP-04)

Em (13), chama atenção o fato de a CV exercer a função de sujeito da sentença encaixada, mas não expressar o papel de agente, mas o de tema, por ser uma construção ergativa, em que o argumento interno do verbo é alçado para a posição de sujeito.

[9] Segundo Baker (1997), a disposição dos elementos na sentença segue preferencialmente uma hierarquia temática que pode ser assim representada: agente/experenciador/causativo > tema > benefactivo/locativo... De acordo com essa hierarquia, o elemento mais alto no VP, o argumento externo, deverá receber papel temático de agente, enquanto o argumento interno, a VP, deverá receber papel temático de tema. De acordo com essa hierarquia, pode-se dizer que o argumento externo, por tender a ser agente, deverá possuir traço [+animado], e o argumento interno, por receber papel de tema, não terá restrição de animacidade.

Voltando ao PL, os dados expostos na Tabela 5 mostram que as estruturas em que o ODA está ligado a um predicativo também favorecem o seu uso quando o antecedente possui traço [+animado] e que o traço [–animado] desfavorece-o totalmente, não havendo uma só ocorrência do PL nesse tipo de estrutura. Quando o traço do antecedente é [–animado], verifica-se que o PL aparece quase sempre em estruturas em que o OD recebe o papel de tema. Em resumo, o PL é favorecido pelo traço [+animado] do antecedente e pelas estruturas complexas, e a CV é favorecida pelo traço [–animado] e pelas estruturas menos complexas.

A fim de verificar, a trajetória do PL nesse dialeto, cruzaram-se os dados referentes à estrutura interna do VP com a faixa etária.

Tabela 6 - Faixa etária *versus* estrutura do SV no dialeto rural afro-brasileiro

ESTRUTURA DO SV	FAIXA 1		FAIXA 2		FAIXA 3	
	CV	PL	CV	PL	CV	PL
	Nº ocorr/%	Nº ocorr/%	Nº ocorr/%	Nº ocorr/%	Nº ocorr/%	Nº ocorr/%
V+OD(SN)	324/**86**	54/14	360/83	72/17	280/**92**	24/8
V+OD(SN)+SP(OI/LOC)	106/**89**	13/11	85/**89**	10/11	72/**94**	5/6
V+OD(SN) + PRED	6/60	4/**40**	7/54	6/**46**	8/89	1/11
$V+_{OD}[_{S}SU_{acc}+V_{[-tempo]}]$	7/41	10/**59**	5/56	4/**44**	2/20	8/**80**
Total	443/85	81/15	457/83	92/17	362/91	38/11

O cruzamento dessas variáveis permite vislumbrar os passos da aquisição do PL no dialeto em estudo. Em estruturas em que o ODA é o sujeito de um verbo da oração encaixada, o PL é favorecido nas três faixas etárias, sobretudo na faixa 3, com 80% das ocorrências (embora sejam apenas 10 dados). O segundo contexto favorecedor do PL é a estrutura em que a variável dependente analisada funciona como sujeito de uma minioração. Essa estrutura também favorece o PL nas três faixas consideradas. Nos demais contextos, o PL é menos frequente, porém a sua frequência de uso aumenta nas faixas constituídas por falantes com menos idade.

Confrontando os resultados das Tabelas 5 e 6, é possível traçar a trajetória do PL nesse dialeto:

(i) inicialmente, o PL ocorre em contextos em que a posição favorece elementos com traço [+animado], como nas estruturas em que o ODA é sujeito de uma oração encaixada, o que ocorre com maior frequência na faixa 3;

(ii) o uso inicial amplia-se nos contextos em que há também complexidade estrutural, mas que não exige necessariamente o traço [+animado] do antecedente, como a minioração entre o ODA é um predicativo;

(iii) por fim, esta variante expande-se para os contextos não complexos, em que o PL só pode ser interpretado como objeto direto e ocorre inclusive com antecedente com traço [–animado].

17.2.4 Referencialidade do objeto direto anafórico

Segundo Cyrino (1997, p.202), o traço referencialidade relacionado com o traço animacidade é bastante relevante no licenciamento do objeto direto que parametriza o PB, pois a CV, no PE, é licenciada quando seu antecedente é determinado/específico, e sua recuperação se dá através de um tópico discursivo ou num contexto pragmático. Já no PB, a CV ocorre mais livremente, com antecedente com traços [+especifíco/referencial] e pode ser recuperada estruturalmente, como se pode ver nos exemplos abaixo:

(14) a. INF: Não, **meu pai** já morava quando eu... /eu num cheguei alcançá meu **pai**$_i$, não, qu'**ele**$_i$ morreu eu tava piquena. É, o povo mataro **ele**$_i$, ele morava em Mariana. (HV-09) [+específico/referencial]

b. INF: Que aqui o ININT mais, mais faço que tinha era... era fazê **um saco de farinha**$_i$ e vende ___$_i$ por dez real (CZ-07) [–específico/referencial]

A Tabela 7 apresenta os resultados desse fator no dialeto aqui analisado.

Tabela 7 - Distribuição das variantes de acordo com a referencialidade do ODA no dialeto rural afro-brasileiro

REFERENCIALIDADE DO ODA	PRONOME LEXICAL			CATEGORIA VAZIA		
	Nº ocorr.	%	p.r.	Nº ocorr.	%	p.r.
[+específico/+referencial]	145/568	26	**.65**	423/568	74	.35
[–específico/+referencial]	54/648	8	.42	594/648	92	**.58**
[–referencial]	14/262	5	.36	248/262	95	**.64**
Total	213/1478	14	—	1265/1478	86	—

A CV é favorecida pelo traço [–referencial] do antecedente com peso relativo .64; enquanto o traço [+referencial] favorece o uso da realização do PL, quando combinado com o traço [+específico] do antecedente, com .65, e favorece a CV, quando combinado com o traço [–especifico], com .58.

17.2.5 Forma verbal

A Tabela 8 apresenta a quantificação dos dados, bem como a relação das formas verbais encontradas.[10]

[10] Alguns tempos verbais foram encontrados em pequeno número, o que levou ao agrupamento de formas semelhantes; o imperativo, por exemplo, foi amalgamado às formas de subjuntivo. Não foi encontrada nenhuma forma de futuro simples, seja do presente, seja do pretérito.

Tabela 8 - Realização do ODA segundo a variável forma do verbo no dialeto rural afro-brasileiro

FORMA VERBAL	PRONOME LEXICAL			CATEGORIA VAZIA		
	Nº ocorr.	%	p.r.	Nº ocorr.	%	p.r.
Pretérito imperfeito	12/113	11	.34	101/113	89	**.66**
Pretérito perfeito	74/425	17	.42	351/425	83	**.58**
Formas compostas	20/204	10	.44	184/204	90	.56
Subjuntivo	10/46	22	.51	36/46	78	.49
Formas nominais	32/239	13	**.57**	207/239	87	.43
Presente do indicativo	65/453	14	**.61**	388/453	86	.39
Total	213/1480	14	—	1267/1480	86	—

A CV é favorecida pelo pretérito perfeito, com .58, assim como pelo pretérito imperfeito, com .66, pois são tempos próprios da narrativa, geralmente ligados ao tópico discursivo. Já o presente do indicativo, como no exemplo em (15), a seguir, remete a ações praticadas com frequência, habituais, e com menor nível de coesão discursiva, favorecendo o uso do PL (.61), garantindo a coesão do enunciado, como se pode ver nos exemplos abaixo.

(15) INF: É. O **cravo**$_i$ esse ano que passou deu o que, dez, doze reais o quilo, já foi uma força né?

DOC: Mas também daqui que o **cravo**$_i$ faça um quilo, né?

INF: Não dá muito trabalho não! Aqui o cara colhe... um cara bom pega uns seis sete quilos de cravo. [...] Você colhe **ele**$_i$, ele tem **aqueles dente**$_i$, aqueles cacho, você quebra **eles**$_i$ sem folha, prá casa, chega em casa, você destala **ele**$_i$, tira **o talo**$_i$, [...] e pega **ele**$_i$ e bota $\underline{\quad}_i$ no só.

O fator *formas nominais*, que inclui as formas do gerúndio e do infinitivo, favorece, segundo Duarte (1986), não só o CL quanto o PL. O PL, no dialeto urbano paulista, é favorecido pelo gerúndio com 23,8%, e o CL é favorecido pelo uso do infinitivo, com 13,5%, pelo fato de sua realização contar com o apoio de uma consoante, estabelecendo o padrão silábico CV, como no exemplo de Duarte (1986, p.15), apresentado em (16).

(16) **Ele**$_i$ veio do Rio só para me ver. Então eu fui ao aeroporto **buscá-lo**$_i$.

No dialeto rural afro-brasileiro, nem mesmo em contextos com o infinitivo essa variante aparece.

17.2.6 Posição da ocorrência em relação ao antecedente

Fixou-se este fator buscando verificar se a distância entre o referente e seu anafórico interferiria na seleção da estratégia do ODA. Na análise desta

variável,[11] observou-se que, quando o antecedente se encontra em outro período, ou seja, o ODA se encontra mais distante do antecedente, como em (16), o PL é favorecido, com .68.

(17) INF: [...] Senhô com vinte lito de cacau, de cacau sequim, qu'é por causo chega aqui que vê tudo maduro, quebra **ele**, fumenta, bota **ele** dento de um saco plástico, (HV-22)

Nos contextos de coordenação, em que o antecedente se encontra na oração imediantenete anterior, a CV é favorecida. Quanto mais próximo estiver o ODA do antecedente, como em (17a), maior é a coesão, portanto a CV pode ser utilizada (com .53) sem comprometimento do enunciado; quanto maior a distância; como em (17b), menor a coesão e o preenchimento da posição é favorecido, com peso relativo igual a .68.

(18) a. INF: Eu vem de manhã, coloco o fêjão$_i$ no fogo e deixa **___**$_i$ aí. (CZ-01)

b. INF: **a menina** evinha da roça, / o sole entrano,/ o cascavéi pegô **ela**,/ a noite inteira ficô ruinzinha... (RC-24)

Nas estruturas de coordenação sindética, independentemente da posição ocupada, é favorecida a CV, com .65, o que demonstra que o elemento conector estabelece coesão entre as orações, permitindo o esvaziamento da posição de objeto direto anafórico.

(19) INF: pegô [**as cerca tudo**]$_i$ e desmantelô**___**$_i$.dano presento pro home... (CZ-12)

Em períodos em que há uma relação de dependência entre as orações, há uma relação de coesão bastante intensa, o que favorece o uso da CV, principalmente quando a CV ocorre na sentença matriz e o antecedente ocorre na adverbial anteposta a ela (.61), como no exemplo a seguir.

INF: [...] Já cavalo, não, quase que num cai não. Mas, se ele tivé de saltá uma cerca$_i$, **ele salta ___**$_i$. (SP-06)

17.3 Condicionamentos sociais

No processamento quantitativo dos dados, o programa VARBRUL selecionou como estatisticamente relevantes as variáveis sociais sexo e comunidade de origem. Os pesos relativos da variável sexo revelaram que as mulheres preferem a CV, enquanto os homens preferem a variante PL. Ambas apresentam o mesmo índice de favorecimento, .58. Assumindo que a CV seja a variante mais antiga nesse dialeto, as mulheres mantêm uma postura conservadora, enquanto os homens inovam mais.

[11] Como foram fixados inúmeros fatores, a tabela ficou extensa; então, preferimos não usar tabelas nos resultados desta variável e expor apenas os seus resultados mais relevantes.

A outra variável social considerada pelo programa foi a variável comunidade de origem. Para a interpretação dos resultados alcançados na análise quantitativa dessa variável, é importante levar em consideração características particulares a cada comunidade:[12]

(i) Cinzento (CZ), de difícil acesso, não mantém contato permanente com outros dialetos; é, portanto, bastante isolada.

(ii) Em Helvécia (HV), os homens, a fim de buscar sustento para a família, saem da comunidade, inclusive para outros estados, entrando em contato com outros dialetos.

(iii) Na região de Rio de Contas (RC), o turismo é uma atividade bastante desenvolvida, o que possibilita a seus membros estar em constante contato com dialetos diferentes;

(iv) Sapé (SP) fica numa região de difícil acesso, e seus membros não demonstram interesse em sair da comunidade; é, portanto, uma comunidade isolada.

Tabela 9 - Distribuição das variantes no dialeto rural afro-brasileiro segundo a comunidade de origem

COMUNIDADE	HV	RC	CZ	SP	TOTAL
	Nº de ocor./%	Nº de ocor./%	Nº de ocor./%	Nº de ocor./%	Nº de ocor./%
SN	49 / 11	70 /16	99 / **20**	57 / **15**	275 / 16
PL	74 / **17**	65 / **15**	40 / 8	34 / 9	213 / 12
CV	307 / 71	306 / 69	359 / **72**	295 / **76**	1267 / 72
Total	430/**25**	441/**25**	498/**28**	386/**22**	1755/100

Os resultados quantitativos revelaram que, nas comunidades mais isoladas, Cinzento e Sapé, a variante CV, considerada a mais antiga, é favorecida apenas em Sapé, com 76%. Em Cinzento, a frequência desta variante reflete a percentagem do cômputo geral dos dados, 72%. Nas demais comunidades, embora desfavorecida, o percentual da CV fica próximo à média de uso dessa variante no dialeto aqui analisado.

O PL é favorecido, embora discretamente, nas comunidades em que há maior contato com outros dialetos, seja pela saída de seus membros em busca de trabalho (Helvécia, com 17%), seja pela interferência do turismo (Rio de Contas, com 15%).

A última variante, o SN, é favorecida nas comunidades em que o PL é menos frequente, em Cinzento e em Sapé, o que lhe confere o caráter de substituto do PL.

Embora a variável faixa etária não tenha sido considerada relevante pelo VARBRUL, é interessante analisar os percentuais referentes a seu uso, a fim de se

[12] Uma análise mais detalhada das características históricas e socioeconômicas de cada comunidade pode ser encontrada no capítulo 2 deste livro.

apresentar mais um argumento para a hipótese de que o contato multilguístico favoreceu a implementação da CV na realização do ODA.

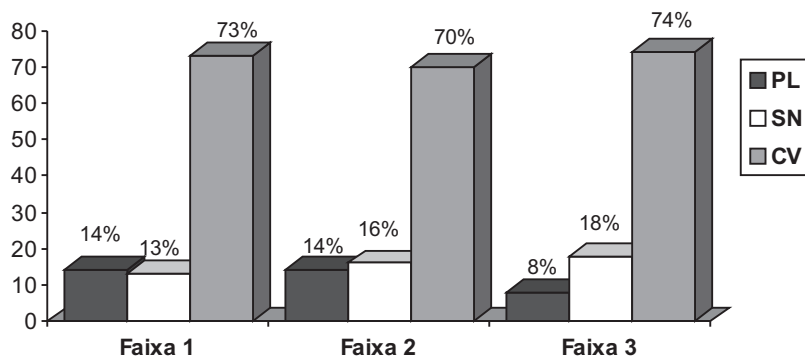



Figura 3 - Distribuição das variantes segundo a faixa etária no dialeto rural afro-brasileiro

Como demonstra a Figura 3, a CV apresenta um comportamento bastante equilibrado em todas as faixas etárias. Se, por um lado, isso não aponta para mudança em curso, por outro, revela que essa variante já está estratificada na comunidade há bastante tempo. Esse comportamento, aliado ao fato de não haver ocorrência de CL, mostra que, na situação de contato, os falantes perderam o CL, uma variante [+marcada] para caso, e passaram a utilizar uma variante neutra, a CV, um comportamento bastante comum em situação de transmissão linguística irregular (cf. capítulo 3 deste livro).

Percebe-se ainda que, no dialeto analisado, o PL foi adquirido posteriormente, tendo em vista a baixa frequência na faixa 3, com 8%, que aumenta nas duas outras faixas para 14%. A frequência do SN diminui à medida que a idade dos falantes diminui. O comportamento desses dois fatores parece apontar para a substituição do SN pelo PL, bem como para a entrada posterior do PL nesse dialeto.

No dialeto urbano paulista (DUARTE, 1986), o uso do PL apresenta o mesmo comportamento encontrado na análise do *corpus* desta pesquisa: é mais frequente entre os jovens e decresce à medida que aumenta a faixa etária.

17.4 Conclusão

Este trabalho, norteado pela hipótese de que a introdução de novas estratégias de realização do ODA se deva ao contato linguístico existente durante o período de colonização, apresentou dados que mostam que a CV foi a variante eleita pelos falantes na situação de contato linguístico, portanto, a variante mais antiga no dialeto rural afro-brasileiro. O equilíbrio no percentual de frequência dessa variante

em todas as faixas etárias — 73% na faixa 1, 70% na faixa 2, e 74% na faixa 3 — confirma tal hipótese. O PL, inicialmente considerada uma estratégia crioulizante, mostrou-se inovadora, tendo em vista sua baixa frequência na faixa 3, com 8%, e seu aumento nas faixas 1 e 2. A sua trajetória pode ser assim descrita:

(i) aparece, na faixa 3 com frequência de 80%, sendo favorecido em construções complexas e com o traço [+animado];

(ii) sua frequência aumenta na faixa 2 em contexto também complexo, que não apresenta restrição de traço de animacidade;

(iii) sem marca de caso e sem especificidade de traço semântico, passa a ocorrer em estruturas não complexas, em que é argumento interno do verbo com papel temático de tema, e sem restrição de animacidade: V+OD(SN) e V+OD+SP.

Não se pode esquecer que a CV é a variante mais realizada, independentemente do traço de animacidade de seu antecedente. Embora tenha se proposto uma trajetória para o PL, sua frequência ainda é bastante baixa.

18 A alternância dativa

Dante Lucchesi
Camila Mello

O processo de transmissão linguística irregular de tipo leve que caracteriza a formação das variedades populares do português brasileiro tem como característica central a erosão na morfologia flexional sem valor informacional, sendo marginais os processos de criação de estruturas originais com valor informacional, que seriam gramaticalizadas no processo de reestruturação da variedade linguística que se forma na situação de contato, como o sistema de tempo, modo e aspecto e os verbos seriais, entre outras estruturas que são típicas das línguas *pidgins* e crioulas, formadas em um processo mais radical de transmissão linguística irregular. Desse modo, os principais processos de variação e mudança induzidos pelo contato entre línguas que se podem identificar hoje no português popular brasileiro (PPB) são os que resultam no amplo quadro de variação no uso das regras de concordância nominal e verbal, bem como na flexão de caso dos pronomes pessoais, entre outros (LUCCHESI, 2003, 2008). Contudo, Lucchesi (2001a) identificou uma estrutura presente no PPB, particularmente nas comunidades rurais de afrodescendentes, que seria um sério candidato a representante de um processo original de reestruturação gramatical em relação à história da língua portuguesa. Trata-se de uma forma de atualização da **estrutura de dativo** que tem sido denominada na literatura de **construção de objeto duplo (COD)** e é encontrada no inglês e no holandês, bem como na maioria das línguas crioulas.

Este capítulo apresenta uma análise sociolinguística da variação na realização da estrutura de dativo que se observa nas comunidades rurais afro-brasileiras isoladas, em que a construção de objeto duplo, exemplificada em (2) abaixo, se encontra em variação com a **construção de dativo preposicionado (CDP)**, exemplificada em (1):

(1) Eu dei o remédio aos meninos.

(2) Eu dei os menino o remédio.

A primeira, CDP, é a variante padrão no português brasileiro (PB) e única opção estrutural disponível nas línguas românicas. A segunda, COD, só é encontrada em variedades populares do PB faladas, não apenas no Estado da Bahia, como também

na Zona da Mata Mineira (SCHER, 1996). Para um falante urbano culto do PB, a COD é uma construção agramatical. Assim, a variação na forma da estrutura de dativo constitui uma dimensão bem radical do quadro de polarização sociolinguística do Brasil, nos termos definidos por Lucchesi (1998, 2001a, 2002a, 2002b, 2006).

Por outro lado, a originalidade da COD *vis-à-vis*, não apenas à história da língua portuguesa, mas ao quadro geral das línguas românicas, descarta qualquer possibilidade de esta construção ser o resultado de um processo interno de mudança ou de uma deriva secular da língua, nos termos de Naro e Scherre (2007). O fato de a COD ser geral entre as línguas crioulas aponta fortemente no sentido de sua presença nas variedades populares do PB estar relacionada historicamente ao massivo contato do português com as línguas indígenas e, muito especialmente, com as línguas africanas. Sendo o resultado do processo de transmissão linguística irregular desencadeado pelo contato entre línguas, essa variação na estrutura de dativo tem uma significação muito especial, pois coloca a influência do contato linguístico na formação do português do Brasil em um novo patamar.

Segundo Lucchesi (2003, p.275 et seq.),

> [...] as características fundamentais do processo de transmissão linguística irregular seriam:
>
> (i) perda, ou variação no uso, de morfologia flexional e palavras gramaticais;
>
> (ii) alteração dos valores dos parâmetros sintáticos em função de valores não marcados, que não implicam, entre outras coisas, em movimentos aparentes na estruturação da sentença; e
>
> (iii) gramaticalização de itens lexicais para preencher as lacunas na estrutura linguística.

Mas, no caso específico da formação do PPB, que se caracteriza como um processo de tipo leve, "os processos originais de reestruturação gramatical seriam raros e periféricos" e as alterações provocadas pelo contato entre línguas na formação do PPB estariam circunscritas aos itens (i) e (ii). E, mesmo nesses casos, não ocorreria uma mudança completa, como no caso dos crioulos típicos, nos quais os morfemas de número e pessoa dos verbos desaparecem, e os sujeitos referenciais nulos deixam de ser licenciados; o que ocorreria seria um amplo processo de variação no emprego desses morfemas e um incremento da realização do sujeito pronominal, configurando um processo em curso de mudança paramétrica (DUARTE, 1993, 1995). Dessa forma, a variação na estrutura de dativo, compreendendo a variante da COD, constitui uma novidade no cenário da formação do PPB, pois instanciaria um processo de gramaticalização de uma estrutura original no universo do desenvolvimento histórico da língua portuguesa.

Este capítulo buscará escrutinar a emergência da COD no PPB como uma resultante do contato do português com as línguas africanas, tendo como base empírica uma análise sociolinguística da variação na estrutura de dativo em quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia. Para cumprir esse objetivo, o texto se estrutura da seguinte maneira. Em sua primeira seção, faz uma descrição das formas através das quais a estrutura de dativo se atualiza nas línguas humanas, com base na teoria da Gramática Gerativa e buscando delimitar o escopo dessa estrutura. A segunda seção trata da estrutura de dativo nas línguas crioulas, entre as quais se observa uma predominância da COD. Com isso se coloca a questão de determinar se essa característica geral das línguas crioulas decorre da emergência de estruturas mais naturais da Gramática Universal (GU) ou da transferência de estruturas gramaticais das línguas do substrato africano. Por fim, esses aportes teóricos serão mobilizados na análise variacionista que se apresenta na seção 18.3. Uma questão perpassa todo o desenvolvimento deste texto. Até que ponto a emergência da COD constitui um processo original de gramaticalização e até que ponto é o resultado de simplificação morfológica. No enfrentamento desse problema, espera-se contribuir não apenas para ampliar os horizontes do conhecimento acerca da história sociolinguística do Brasil, mas também para aprofundar a compreensão sobre como os processos de contato linguístico massivo, radical e abrupto afetam a estrutura gramatical da língua.

18.1 A alternância dativa: configurações e escopo

Nas construções de dativo, a estrutura argumental do verbo compreende, além do argumento externo sujeito, dois argumentos internos complementos para expressar a transferência de algo, o TEMA, entre um AGENTE e um BENEFICIÁRIO. No latim, essas funções eram expressas por uma marcação morfológica de caso, como se pode ver no exemplo (3) abaixo, em que o TEMA, a coisa transferida, recebe o **caso acusativo**, enquanto o BENEFICIÁRIO recebe o **caso dativo**. No desenvolvimento do latim para as línguas românicas, essa marcação morfológica de caso se perdeu, sendo substituída pelo emprego da preposição *a*, para marcar o caso dativo do constituinte denominado **objeto indireto (OI)** pela tradição gramatical, enquanto o caso acusativo é atribuído pelo verbo ao constituinte denominado tradicionalmente **objeto direto (OD)**, como se pode ver nos exemplos (4), (5) e (6), do italiano, do espanhol e do português, respectivamente. Nessas línguas, a ordem normal é OD-OI, podendo haver a inversão OI-OD em contextos específicos.

(3) Pulcr**ae** puel**ae** rosa**m** dedi.

Bonita-**dat** menina-**dat** rosa-**acus** dei

(4) **Ho datto** *la rosa* **a** la bella bambina.[1]

(5) **Di** *la rosa* **a** la niña guapa.

(6) **Dei** *a rosa* **à** menina bonita.

Já línguas da família germânica, como o inglês e o holandês, apresentam uma estrutura alternativa, em que um grupo formado pelos dois argumentos internos, na ordem OI-OD, recebe caso conjuntamente do verbo, configurando uma estrutura fechada, semelhante a uma *pequena oração* (ing. *small clause*). Assim, no que tem sido denominado de **alternância dativa**, línguas, como o inglês, dispõem de duas construções para realizar a estrutura de dativo, a construção de dativo preposicionado (CDP), exemplificada em (7), e a construção de objeto duplo (COD), exemplificada em 8.

(7) I **gave** the *rose* **to** the beautiful girl.

(8) I **gave** the beautiful girl *the rose*.

No âmbito da teoria da gramática, predomina a visão de que a CDP seria a construção básica de dativo, sendo a COD uma construção derivada da primeira (LARSON, 1988).[2] Por outro lado, essa derivação não se aplicaria em todos os casos em que ocorre um OI regido pela preposição *to* ou *for*, como se pode ver no contraste dos exemplos (9) e (10) com os exemplos (11) e (12), retirados de Grimshaw (1988, p.114):

(9) We gave a book to the child.
We gave the child a book.

(10) We'll bake a cake for you.
We'll bake you a cake.

(11) We carried a book to the child.
*We carried the child a book.

(12) They'll fix the radiator for you.
*They'll fix you the radiator.

Com isso, coloca-se a questão dos limites da estrutura dativa *stricto sensu*, pois nem sempre o constituinte que se classifica como OI expressa um papel temático de BENEFICIÁRIO. Há construções em que esse constituinte expressa um papel temático de META; nesses casos, a alternância dativa não é possível, como

[1] Nestes exemplos das línguas românicas, bem como nos exemplos do inglês a seguir, os atribuidores de caso estão em negrito, o OD está em itálico e o OI está sublinhado.

[2] Para uma discussão sobre as formalizações analíticas do fenômeno no âmbito da Gramática Gerativa, veja-se Scher (1996).

exemplificado em (11). Já quando a estrutura expressa uma relação de *finalidade* (do tipo 'fazer algo para alguém'), a possibilidade de alternância dativa é variável, como se pode ver no cotejo entre os exemplos (10) e (12). Por isso, Grimshaw (1988) afirma que, para haver alternância dativa, o OI tem de receber o papel temático de BENEFICIÁRIO, implicando uma relação de **transferência de posse**, o que ocorre em (10), mas não ocorre em (12). Porém, as construções dativas com verbos como *donate*, *contribute* e *distribute* não admitem a alternância, mesmo havendo a transferência de posse, como na frase exemplificada em (13), o que deixa ainda em aberto a questão do escopo da transferência dativa em inglês.[3]

(13) I distributed apples to the children. [4]

*I distributed the children apples.

Morais, Ribeiro e Ferreira (no prelo) também se colocaram o problema do escopo da estrutura de dativo, mas no âmbito das línguas românicas, particularmente no espanhol e no português. No caso dessas línguas, o problema em delimitar a estrutura de dativo decorreria "do fato de que os conteúdos cobertos pelo caso dativo do latim passaram e ser expressos pelas preposições *a* e *para* nas línguas românicas". Contudo, com base em estudos sobre o fenômeno no espanhol moderno (HERNANZ; BRUCART, 1987; GATIÉRREZ ORDÓÑEZ, 1999), esses autores observaram que há contextos que admitem a variação entre as duas preposições, como exemplificado em (14), e há outros que não admitem a variação, como se pode ver em (15).

(14) a. Compró una chaqueta a Juan.

b. Compró una chaqueta para Juan.

c. Hizo una chaqueta al niño.

d. Hizo una chaqueta para el niño.

(15) a. Les preguntó la lección a los alumnos.

b. *Les preguntó la lección para los alumnos.

c. *Pedro trabaja a sus hijos.

d. Pedro trabaja para sus hijos.

A impossibilidade de emprego da preposição *a* em (15)c. deve-se ao fato de o OI realizar "a função semântica de finalidade, valor semântico exclusivo da preposição *para*"; ao passo que "a função semântica beneficiário é característica da

[3] Scher (1996, p.24) chama a atenção para o fato de os três verbos em questão serem de origem latina, o que pode apontar para uma interferência significativa de um aspecto sócio-histórico na estruturação gramatical da língua.

[4] Exemplo retirado de Larson (1988, p.591).

preposição *a*", o que interditaria o uso de *para* em (15)b. Outro parâmetro para delimitar o fenômeno em espanhol seria o redobro do clítico. Só as construções com *a* permitem o redobro do clítico:

(16) a. Le envió una postal a Pepe.

b. *Le envió una postal para Pepe.

Os autores chegam à conclusão de que, no espanhol, só as verdadeiras construções de dativo permitem o redobro do clítico e podem ser introduzidas pela preposição *a*.

No âmbito da língua portuguesa, Morais, Ribeiro e Ferreira (no prelo) identificaram diferenças significativas na forma como o português europeu (PE) e o português brasileiro (PB) realizam a estrutura dativa, como se pode ver nos exemplos abaixo:

(17) a. O José enviou uma carta **à Maria**/enviou-**lhe** uma carta. (PE/PB)

b. O José enviou uma carta **para a Maria/à Bélgica**. (PE/PB**)**

c. O José atribuiu uma tarefa **aos alunos**/atribuiu-**lhes** uma tarefa. (PE/PB)

d. O José atribuiu uma tarefa **para os alunos**. (*PE/PB)

e. O José devolveu o dinheiro **aos clientes**/devolveu-**lhes** o dinheiro. (PE/PB)

f. O José devolveu o dinheiro **para os clientes**. (*PE/PB)

g. Maria comprou o livro **ao amigo**/comprou-**lhe** o livro. (PE/*PB)

h. Maria comprou o livro **para o amigo**. (PE/PB)

(18) a. Morreu-**lhe** o filho mais velho. (*PB/ PE)

b. Falta-**lhe** uma boa esposa. (*PB/ PE)

(19) a. O José deu o livro **a ela / para ela**. (PB/*PE)

b. O José deu-**lhe** o livro **a ela**.[5] (*PB/ PE)

Como se pode ver em (17), a preposição *a* é muito mais produtiva no PE, pois no PB a preposição *para* avançou bastante sobre os contextos outrora privativos daquela. O emprego do clítico com flexão de dativo *lhe* foi igualmente restringido no PB, como se pode ver em (18) e (19), com o redobro do clítico só sendo possível no PE.[6] Diante desse quadro, Morais, Ribeiro e Ferreira (no prelo) constataram uma inovação na forma como o PB passou a atualizar a estrutura de dativo, em oposição ao PE, no qual se conservaram as construções encontradas nos estados

---

[5] Neste caso, o OI tem um valor contrastivo.

[6] Essas assertivas dizem respeito ao uso natural da língua em sua modalidade oral. No Brasil, a tradição gramatical tem difundido as formas naturais no PE, mas ausentes no PB, fazendo com que algumas dessas construções que não fazem parte da gramática natural dos brasileiros possam ser encontradas em seus registros de fala formal ou retórica e na língua escrita.

anteriores da língua. A inovação do PB se caracteriza, então, pela "perda dos pronomes clíticos dativos de 3ª pessoa e pela perda da preposição *a*, na sua função gramatical de marcar morfologicamente o caso dativo". Os autores chegam à conclusão de que "a sintaxe do dativo, no PB, não se configura como resultado de deriva, de continuidade de propriedades sintáticas da língua mãe". Coloca-se a questão de saber:

> [...] até onde a situação de aquisição imperfeita do português, no sentido de Kroch (2005) e Lucchesi (2000, 2003), típica do contato contínuo e intenso entre línguas — como aconteceu, durante todo o período colonial, entre o português e as línguas indígenas e africanas —, pode estar subjacente às mudanças que ocorreram no PB em relação às construções dativas (MORAIS; RIBEIRO; FERREIRA, no prelo).

As mudanças no PB configuram um processo de simplificação morfológica, com a perda das marcas exclusivas de dativo, acompanhada da expansão do uso de uma preposição multifuncional, a preposição *para*, que é empregada para expressar os mais diversos papéis temáticos: LOCATIVO, BENEFICIÁRIO, META, etc. Por outro lado, ratificando mais uma vez o quadro de polarização sociolinguística do Brasil (LUCCHESI, 1998, 2001a, 2002a, 2002b, 2006), paralelamente a essas mudanças que afetam todas as variedades do PB, observa-se a emergência das construções de objeto duplo em algumas variedades populares do PB, configurando-se o fenômeno da alternância dativa. A radicalidade desse processo de mudança, que gerou construções absolutamente estranhas aos falantes cultos das grandes cidades brasileiras, coloca de forma incontestável a relação histórica da emergência da alternância dativa com as situações de contato entre línguas que marcam a formação das variedades populares do PB. Desse modo, ganha relevo o estudo de Scher (1996), que buscou descrever as propriedades da alternância dativa em uma variedade do PPB.

Analisando o fenômeno no que chamou de *dialeto da Zona da Mata Mineira*,[7] Scher (1996, p.30-31) observou que a COD pode exibir também a ordem OD-OI, e não somente a ordem OI-OD, como exemplificado em (20)c. e (20)b., respectivamente. Essa inversão da ordem na COD também foi observada na análise sociolinguística das comunidades afro-brasileiras isoladas e pode ser atribuída a uma possível transferência do substrato africano (cf. seção 18.3, abaixo).

(20) a. Eu emprestei o carro ao Pedro.

b. Eu emprestei o Pedro o carro.

c. Eu emprestei o carro o Pedro.

---

[7] Infelizmente, a autora não dá qualquer informação sobre a distribuição sociolinguística do fenômeno (cf. p.13, nota 6).

A autora (1996, p.34-38) também se colocou o problema do escopo da alternância dativa no dialeto que estudou e observou que a COD não é possível quando o OI só pode ser introduzido pela preposição *para* (cf. exemplos em (21)). Já quando se observa a variação entre *para* e *a*, a possibilidade de alternância é variável, como se pode ver nos exemplos em (22), em que a alternância dativa não é possível em a. e b., mas o é em c. e d.

(21) a. buscar flores para a namorada

b. buscar as crianças para a mãe

c. cantar (uma canção) para os amigos

d. comprar um presente para a mãe

(22) a. deixar muitos bens para/a os filhos

b. descrever o evento para/a os presentes

c. anunciar a decisão (para/a) os filhos

d. dar (um presente) (para/a) o pai

Segundo Scher (p.37), essa variabilidade poderia ser explicada pelo fato de verbos como *descrever* e *insinuar* poderem figurar "em sentenças sem a presença obrigatória do OI (*Ele descreveu o evento muito bem, A Joana adora insinuar bobagens*, etc.)", concluindo que "na realidade [esses] verbos não preveem, em sua grade temática, a realização de um argumento *Alvo*". A fragilidade desse argumento está no fato de, nas frases citadas, a referência do OI estar implícita, o que pode ocorrer com qualquer verbo que desencadeia a estrutura de dativo. Portanto, o estudo de Scher não chega a uma conclusão sobre os limites do fenômeno no dialeto da Zona da Mata Mineira, de modo que, assim como no inglês, não se dispõe de parâmetros que possam delimitar o escopo do fenômeno da alternância dativa.

Por outro lado, Scher (1996) chamou a atenção para diferenças estruturais que distinguem o fenômeno no inglês e no PPB. Uma dessas diferenças seria a possibilidade de construção passiva com o OI ocupando a posição de sujeito, que não existe no PPB, mas existe no inglês, como se pode ver no exemplo (23)d.:

(23) a. I gave a book to Mary.

b. I gave Mary a book.

c. A book was given to Mary.

d. Mary was given a book.

Outra diferença é a possibilidade de COD com o OD não realizado foneticamente, que ocorre no dialeto analisado por Scher, mas não ocorre no inglês. Essa possibilidade de alternância dativa com o OD nulo também foi observada nas comunidades rurais afro-brasileiras isoladas (cf. seção 18.3 abaixo) e enquadra-se

perfeitamente no contexto de generalização do objeto nulo no PB (DUARTE, 1989 e o capítulo anterior).

Diante dessas e de outras diferenças, Scher conclui que, as construções encontradas no dialeto que estudou e no inglês “não representam o mesmo fenômeno sintático” (p.43). Tal afirmação é bastante discutível fora do quadro teórico da autora, mas é razoável pensar que o mesmo fenômeno se apresenta de formas diferenciadas nas duas variedades linguísticas; o que de resto é muito evidente considerando o universo mais amplo das línguas em que se observa a alternância dativa (cf. seção 2, a seguir). Há, inclusive, um paralelo interessante na forma como o parâmetro do sujeito nulo se apresenta em línguas como o inglês e o francês, de um lado, e nas línguas crioulas, de outro. Nessas últimas, a impossibilidade de um sujeito referencial nulo não implica necessariamente a existência de um sujeito expletivo sem valor referencial, como nas primeiras (cf. capítulo 6 deste livro). Portanto, fora de uma visão muito estreita da teoria da gramática, pode-se assumir que as propriedades da gramática se manifestam de maneira variável nas línguas humanas.

Além disso, tal encaminhamento desvia a abordagem da questão mais relevante a se destacar, que é a presença no PPB de uma construção original no universo das línguas românicas e que, por conseguinte, não pode ser pensada em termos de uma evolução interna da língua portuguesa ou de uma alegada deriva secular românica, nos termos de Naro e Scherre (2007). O fato de a COD ser geral nas línguas crioulas, mesmo naquelas derivadas de línguas românicas, não deixa margem à dúvida de que o fenômeno foi desencadeado nas situações de contato do português, particularmente, com as línguas africanas. A análise da estrutura de dativo nas línguas crioulas contida na próxima seção pretende reunir evidências empíricas que possibilitem uma melhor compreensão da presença da alternância dativa nas variedades populares do PB, particularmente no português afro-brasileiro.

18. 2 A estrutura de dativo nas línguas crioulas: universais *versus* substrato

Em um amplo estudo sobre a estrutura de dativo nas línguas *pidgins* e crioulas, Bruyn, Muysken e Verrips (1999) observaram que a COD é geral nessas línguas, tanto naquelas línguas lexificadoras, como o inglês e o holandês, que oferecem essa possibilidade — cf. exemplos (24) e (25) abaixo, do tok pisin e do negerhollands, respectivamente —, quanto naquelas derivadas de línguas que não possuem tal construção, como o santomense, de base lexical portuguesa, e o

haitiano, de base lexical francesa — cf. exemplos (26) e (27), respectivamente.[8] Dessa forma, a COD pode ser relacionada historicamente às situações de contato entre línguas massivo, radical e abrupto, sendo definida como uma propensão estrutural do que Lucchesi (2003, 2008) tem denominado **transmissão linguística irregular**.

(24) Mi soim yu banara bilong mi.
1p mostrar 2p arco PREP 1p
[eu mostrei você de eu]
Eu lhe mostrei meu arco.

(25) Ham a gi di man si gout
3p PASS dar DET homem POSS dinheiro
[Ele deu o homem seu dinheiro]
Ele deu seu dinheiro ao homem.

(26) E da bo pichi.
3p dar 2p peixe
[Ele deu você peixe]
Ele deu o peixe pra você.

(27) Li rakonete papa-li istwa sa-a.
3p contar pai-3p história DET
[Ele contou seu pai esta história]
Ele contou esta história pro pai dele.

Ao lado da COD, as línguas crioulas costumam exibir outras duas possibilidades estruturais de marcar o dativo: a CDP, e uma construção em que uma forma verbal é empregada com a função da preposição de dativo. Nesse caso, a perda da preposição da língua lexificadora é compensada com a gramaticalização de um verbo (geralmente o verbo *dar*), que passa a fazer as vezes de preposição, ao lado de seu emprego original como verbo. Esse fenômeno é muito geral nas línguas crioulas, e esses verbos são denominados **verbos seriais**. Assim, a **construção serial de dativo** (CSD) nada mais é do que um sucedâneo da CDP, podendo-se assumir que elas são estruturalmente equivalentes. As três possibilidades estruturais podem ser encontradas em uma mesma língua crioula, como exemplificado abaixo com frases do fa d'ambu, um crioulo de base lexical portuguesa da Ilha de Annobon (BRUYN; MUYSKEN; VERRIPS, 1999, p.330-331):

[8] Os exemplos (24-27) são de Bruyn, Muysken e Verrips (1999, p.330).

(28) Malía da pe-d'eli tabaku. (COD)
Maria dar pai-POSS tabaco
[Maria deu pai dela tabaco.]
Maria deu tabaco pro pai dela.

(29) No sa fe wan xadyi pa non-tudu. (CDP)
4p PROG fazer DET casa PREP 4p-todos
Nós estamos fazendo uma casa para nós todos.

(30) Amu da wan kuzu da bo. (CSD)
1p dar DET coisa dar 2p
[Eu dei uma coisa dá você]
Eu dei uma coisa pra você.

Contudo, das três possibilidades, a COD seria a construção mais geral entre as línguas crioulas. Com base em um estudo comparativo entre quase vinte línguas crioulas, Bruyn, Muysken e Verrips (1999, p.338 et seq.) encontraram CODs em praticamente todas (à exceção do malaio português). Já as CDPs e as CSDs se distribuem com uma frequência bem menor no universo observado, sendo a primeira um pouco mais frequente do que a segunda. Os resultados dessa pesquisa são sumarizados no Quadro 1.

LÍNGUAS CRIOULAS	COD	CDP	CSD
Sranan	+	+	+
Saramacan	+	(-)	+
Crioulo Jamaicano	+	(+)	-
Tok Pisin	+	+	-
Negerhollands	+	+	(-)
Berbice Dutch	+	-	+
Afrikaans	(+)	+	-
Crioulo Haitiano	+	-	+
Guyanais	+	-	-
St. Lucian	+	-	-
Louisiana Creole	+	(+)	-
Seychelles Creole	+	(+)	-
Papiamentu	+	-	-
Palenquero	+	-	-
Guinea-Bissau Kriyol	+	(-)	-
Fa d'Ambu	+	(-)	+
Principense	+	-	+
Malayio-Portuguese	-	+	-
Sri Lanka Portuguese	(+)	(+)	-

Quadro 1 - Distribuição de CODs, CDPs e CSDs em algumas línguas crioulas
Fonte: Bruyn, Muysken e Verrips, 1999, p.338.

Assumindo que a COD seria a estrutura básica nas línguas crioulas, coloca-se a questão de definir as suas origens nos termos da polarização que tem marcado o desenvolvimento da crioulística nas últimas décadas: os universais da linguagem que atuam na crioulização *versus* a transferência do substrato na expansão do *pidgin*.[9] Para os crioulos da região do Caribe, haveria indícios de que a COD poderia ser o produto da transferência de estruturas gramaticais das línguas do substrato, pois as CODs ocorrem em quase todas as línguas do Oeste Africano, exceto nas línguas mandê (KOOPMAN, 1986, p.235), e, na família banto, apresenta-se como praticamente a única possibilidade estrutural na língua quicongo, como exemplificado em (31) e (32) abaixo (LUMWANU, 1973, p.137). A língua ewe, da família kwa, exibe uma curiosidade: a COD só admite a ordem OD-OI — como exemplificado em (33), extraído de Bole-Richard (1983, p.227) —, e não a ordem OI-OD encontrada na imensa maioria das línguas que têm COD.

(31) Ganá mwa:na mamba.
dar criança água
Dê água pra criança.

(32) Ganá Nkuká mwana.
dar Nkuká criança
Dê a criança pra Nkuká.

(33) Na èlà dèví-ó-cá.
Dar carne criança-PL-também
Dê a carne pras crianças também.

A dificuldade para identificar historicamente as línguas de substrato que estiveram na base da formação de cada língua crioula em particular tem comprometido a validade heurística da hipótese do substrato.[10] Além disso, no caso particular da COD, a sua alegada presença nas línguas das famílias kwa e banto não pode explicar a sua presença em línguas crioulas com outros substratos, como as dos oceanos Índico e Pacífico. A presença generalizada da COD nas línguas crioulas de diferentes regiões do planeta parece militar em favor de hipóteses universalistas. Por outro lado, a presença da CSD em vários crioulos do Caribe, como o saramacan, o sranan e o haitiano, bem como nos crioulos portugueses do Golfo da Guiné (santomense, principense e fa d'ambu), parece estar relacionada com uma forte influência do substrato (BRUYN; MUYSKEN; VERRIPS, 1999, p.362).

Considerando-se a hipótese universalista, a presença generalizada da COD nas línguas crioulas implicaria que essa seria uma opção não marcada da GU. Os

[9] Cf. capítulo 3 deste livro.

[10] Idem.

resultados apresentados por Bruyn, Muysken e Verrips (1999, p.356-361) de estudos de aquisição do inglês e do holandês como língua materna tendem a favorecer essa hipótese, na medida em que, "no holandês da criança, CODs aparecem antes das CDPs, assim como no inglês, ao menos quando mensurado em termos de primeira ocorrência na fala espontânea".[11] Contudo, o estudo comparativo entre o haitiano e a aquisição do francês como língua materna, apresentado por esses autores (1999, p.362-366), parece negar a hipótese forte do **parâmetro não marcado**.

Se a emergência da COD em crioulos derivados das línguas românicas, como o haitiano, for o resultado do processo de aquisição da língua materna, e não de uma transferência do substrato, que teria introduzido ocorrências de COD na variedade de francês falada como segunda língua pelos escravos africanos que forneceu o *input* para aquisição da língua materna no processo de crioulização, as crianças deveriam produzir estruturas de COD independentemente do *input* que recebessem. Porém, os resultados apresentados por Bruyn, Muysken e Verrips (1999, p.361-362) para a aquisição do francês como língua materna não contêm ocorrências de COD. Esses autores, então, recorrem à ideia de uma relatividade do conceito de parâmetro não marcado, no sentido de que o parâmetro não marcado seria mais fácil de adquirir, não obstante também carecer de evidência positiva para ser desencadeado na fala das crianças, e defendem que a presença de COD no haitiano poderia ser explicada como uma análise abdutiva[12] da estrutura imperativa do francês com um clítico pós-verbal com a função de OI, como na frase: *donne-moi du pain* ('dê-me o pão'). Tal explicação parece um tanto quanto *ad hoc* e bastante específica, pois, na melhor das hipóteses, poderia explicar a presença de COD no haitiano, mas não a sua presença generalizada nas línguas crioulas.

Uma visão alternativa seria a de Bickerton (1999) de **valor *default***, segundo a qual o valor *default* é desencadeado no processo de crioulização, caso não haja qualquer evidência em outra direção. Ou seja, as crianças no processo de crioulização optariam pela COD, nos casos em que não encontrassem no *input* dados robustos de uma morfologia de dativo, fosse ela uma preposição, fosse uma serialização verbal. Tal visão é negada pelo fato de línguas crioulas exibirem tanto a COD quanto a CSD, como é o caso do haitiano e dos crioulos portugueses do Golfo da Guiné, ou até as três possibilidades estruturais, incluindo a CDP, como é o caso do sranan e do fa d'ambu. Aliás, entra em contradição também com opcionalidade estrutural entre a COD e a CDP do inglês e do holandês. O fato de haver línguas que oferecem essas duas opções estruturais e línguas, como as românicas, que só possuem a CDP aponta para a necessidade de evidência positiva para o

[11] Traduzido do original.

[12] Para uma visão acerca da abdução na aquisição da língua materna, vejam-se as seções 4.1.2. e 4.1.3. do capítulo 4 deste livro.

desencadeamento da COD no processo de aquisição da língua materna. Mas isso entra em contradição com a assunção de que a COD seria geral nas línguas crioulas. Além do que, mantém sem solução o problema de explicar a emergência da COD nos crioulos cuja língua lexificadora é uma língua românica, que não tem COD. Com isso, retorna-se à hipótese do substrato, que, como já se viu, também não é compatível com a ideia de um caráter universal da COD.

De qualquer forma, a fragilidade fonética da preposição de dativo *a* das línguas românicas contribuiria para o colapso da CDP no processo de crioulização, possibilitando a emergência ou da COD ou da CSD através de um processo de **relexificação**, conforme proposto por Lumsden (1999) entre outros. A manutenção da CDP no processo de crioulização parece estar associada não apenas à robustez fônica da preposição de dativo, como também à existência de uma preposição multifuncional que desempenhe uma função locativa juntamente com a função de dativo, como *to* no inglês e *para* no português brasileiro, o que é relevante para o caso em foco.

A situação da estrutura de dativo nas línguas crioulas pode ser então sumarizada da seguinte forma. A COD parece ser a alternativa preferencial nos processos de crioulização/pidginização, não obstante haja uma contradição com o fato de que, em princípio, a aquisição da COD dependa de evidência positiva; contradição essa que ainda permanece sem solução no estágio atual da teoria da GU e da aquisição da linguagem. Dessa forma, a COD emergiria nos processos de transmissão linguística irregular em que não houvesse uma morfologia forte de dativo. Essa morfologia entraria no *input* da crioulização, ou na forma de uma preposição foneticamente robusta e multifuncional da língua do superstrato, ou na forma de uma serialização verbal proveniente de um processo de relexificação a partir da gramática das línguas do substrato. Contudo, um processo de variação com a concorrência de gramáticas distintas na comunidade crioulófona emergente[13] pode resultar na adoção das duas opções, dando ensejo a línguas crioulas que, ao lado da estrutura básica COD, exibem as duas construções, tanto a CDP, quanto a CSD. Ou seja, a COD seria a forma não marcada que emergiria na ausência de uma morfologia robusta de dativo no *input* da crioulização; a CSD seria desencadeada a partir de uma morfologia de dativo disponível no *input* da crioulização a partir da transferência gramatical das línguas do substrato, via relexificação, no sentido de Lumsden (1999) e outros; a CDP provém da morfologia de dativo da língua do superstrato (nesse sentido, a sua presença em algumas línguas crioulas poderia ser um fenômeno recente, decorrente da **descrioulização**).

A análise variacionista que se apresenta na próxima seção tem por base essa sistematização, bem como os aspectos teóricos reunidos na seção anterior.

[13] Cf. seção 4.2. do capítulo 4 deste livro.

18.3 A alternância dativa nas comunidades rurais afro-brasileiras isoladas: uma análise sociolinguística

A análise da variação na forma da estrutura de dativo no português afro-brasileiro, com arcabouço teórico-metodológico da Sociolinguística Quantitativa (LABOV, 1972a, 1982, 1994), teve como universo de observação as quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia estudadas pelo Projeto *Vertentes*: Cinzento, Helvécia, Rio de Contas e Sapé.[14] A amostra de fala vernácula analisada foi constituída por 48 entrevistas de tipo sociolinguístico, sendo doze de cada comunidade, e estruturada segundo as variáveis **faixa etária** (faixa 1, de 20 a 40 anos; faixa 2, de 41 a 60 anos; e faixa 3, acima de 60 anos) e **sexo**. Na análise do encaixamento social, foram consideradas ainda as seguintes variáveis não estratificadas na amostra: **estada fora da comunidade** (ausência ou não da comunidade por pelo menos seis meses) e **nível de escolaridade** (analfabeto ou semianalfabeto). O processamento quantitativo dos dados foi feito através do pacote de programas VARBRUL (PINTZUK, 1988; SCHERRE; NARO, 2003; TAGLIAMONTE, 2006).[15]

Na definição da variável dependente, foram estabelecidos os seguintes valores:[16]

1. construção com dativo preposicionado (CDP)

$V + SN_{OD} + SP_{OI}$

Ex.: *ela dá comida pas criança.* (CZ-01)

$V + (SN_{OD}) + SP_{OI}$

Ex.: *deu p'os dois...* (RC-07)

$V + SP_{OI} + SN_{OD}$

Ex.: *escreveu pra nós uma carta.* (RC-21)

2. construções de objeto duplo (COD)

$V + SN_{OI} + SN_{OD}$

Ex.: *ele vendia compade Jacó porco gordo.* (HV-20)

$V + SN_{OI} + (SN_{OD})$

Ex.: *deu meu irmão*. (SP-05)

---

[14] Essas comunidades são apresentadas no capítulo 2 deste livro.

[15] A metodologia empregada na pesquisa é descrita no capítulo 5 deste livro.

[16] A partir de agora, os exemplos são constituídos por ocorrências retiradas da amostra de fala analisada, sendo indicado entre parênteses, o número do informante e a comunidade a que ele pertence, através das siglas: CZ para Cinzento, HV para Helvécia, RC para Rio de Contas e SP para Sapé.

$V + SN_{OD} + SN_{OI}$

Ex: *'cê manda pedi um empresti ele* (HV-20)

Como se pode ver nas ocorrências retiradas da amostra de fala analisada, as duas construções apresentam uma variação na ordem, OD-OI ou OI-OD, e a possibilidade de um OD nulo. A base de dados da análise variacionista foi constituída, em um primeiro momento, por 358 ocorrências de estruturas dativas encontradas com o levantamento exaustivo nas mais de 40 horas de fala vernácula analisadas. Nessa recolha, foram excluídas as ocorrências em que uma eventual crase da preposição com a vogal final do verbo ou da vogal inicial da palavra seguinte, exemplificadas nos exemplos (34) e (35), respectivamente, levasse à neutralização da variável.

(34) a gente chega e progunt**a** o dono da casa se ele aceita (CZ-04)

(35) ela dá bença **a**s madrinha (CZ-02)

No conjunto das 358 ocorrências que constitui a base de dados inicial desta análise, 83% (296 ocorrências) eram da CDP, enquanto apenas 17% eram da COD (62 ocorrências). O OD não foi realizado em 123 ocorrências de CDP, o que corresponde a uma frequência de um pouco mais de 40% de OD nulo nessas construções. No universo restante de 173 ocorrências, 145 apresentaram a ordem $SN_{OD} + SP_{OI}$ (84% do total), de modo que a ordem inversa $V + SP_{OI} + SN_{OD}$ ocorreu com uma frequência de pouco mais de 15%. No universo das COD, o OD nulo também correspondeu a 40% do total de ocorrências (25 em um total de 62). Já a inversão $V + SN_{OD} + SN_{OI}$ foi de 41% do total de ocorrências com o OD realizado (15 contra 22 da ordem $V + SN_{OI} + SN_{OD}$). Esse dado é muito significativo, considerando-se que essa ordem invertida é a única opção disponível na língua ewe (cf. seção anterior), uma das línguas com uma forte presença no substrato africano do Brasil (cf. capítulo 1 deste livro).[17]

Os primeiros resultados da análise podem, então, ser sumarizados da seguinte maneira: um OD nulo ocorreu em quase a metade das ocorrências, tanto da CDP quanto da COD, o que se ajusta à alta frequência de OD nulo no PB. Já a inversão da ordem básica foi baixa na CDP e bem mais elevada na COD, constituindo nesta última uma forte evidência em favor de um processo histórico de transferência gramatical do substrato africano.

[17] Os ewes, ou os falantes do grupo ewe-fon em geral, receberam no Brasil a denominação genérica de *jejes* e foram, juntamente com os iorubás, os grupos etnolinguísticos mais representativos no Brasil da família kwa. Os testemunhos históricos disponíveis indicam a sua presença nos estados da Bahia, Pernambuco, Maranhão, Minas Gerais e Rio de Janeiro. Sua importação teria se iniciado no século XVII, acentuando-se nos séculos XVIII e XIX (CASTRO, 2002, p.43-54).

18.3.1 Condicionamento linguístico da alternância dativa

No plano estrutural, um aspecto crucial do fenômeno em foco é a natureza do verbo (cf. seção 18.1 acima), até mesmo no sentido de estabelecer o seu escopo e *a fortiori* definir a variável dependente desta análise. No sentido mais lato, a estrutura de dativo pode-se atualizar com verbos que foram classificados nesta análise da seguinte maneira: **benefativos**, **leves**, ***discendi***, ***faciendi***, **existenciais**, **de condução** e **de transferência**. Os **verbos benefativos**, que seriam os verbos da estrutura dativa *stricto sensu*, implicam a transferência de algo para outrem; este é o caso dos seguintes verbos: *dar, pedir, doar, mostrar, encomendar, devolver, emprestar, vender, entregar, ceder, passar, pagar*, etc. Os **verbos leves**, que formam uma lexia composta com o seu OD, recobrem uma ampla área semântica, como se pode ver nos seguintes exemplos: *dar apoio, dar uma mão, fazer promessa, dar trabalho, fazer uma leitura, dar prejuízo, dar aula, dar aviso, dar má criação, dar explicação, dar um agrado, dar saúde, dar problema, dar assistência,* (*não*) *ter consideração, dar socorro, dar emprego, fazer mal, fazer bem, fazer medo,* etc. Os **verbos discendi** remetem a comunicação verbal de um conteúdo a outrem, como no caso dos verbos: *dizer, falar, contar, perguntar, ensinar, agradecer, escrever, explicar, avisar, alertar*, etc. Os **verbos faciendi** indicam uma ação feita para outrem, com em: *fazer algo para alguém, comprar algo para alguém, preparar algo para alguém, pagar algo para alguém,*[18] *olhar* (*cuidar*) *para, arar para, resolver para, pegar para, liberar, receber, tirar, arrumar, copiar, procurar, colocar, costurar, arranjar, tomar conta,* etc. Os **verbos existenciais** denotam a existência ou ausência de algo ou alguém que cabe a outrem, ou o afeta de alguma maneira, de que são exemplos: *ter algo* (*para alguém*), *faltar algo* (*para alguém*), etc. Os **verbos de condução** referem-se ao transporte ou à remessa de algo para outrem, como no caso dos verbos: *levar, encaminhar, conduzir, despachar, trazer, botar*, etc. Por fim, os **verbos de transferência** indicam a outorga de algo para outrem, como no caso de: *deixar, ficar, caber, distribuir*, etc. Abaixo, são apresentadas ocorrências da amostra de fala analisada para exemplificar cada um desses tipos:

1. benefativos

(36) A coisa **deu** *um terreno* pra ele (RC-09)[19]

(37) **deu** o japonês *vinte mil* (HV-20)

[18] Deve-se fazer a diferença entre (i) e (ii) abaixo. Em (i) a construção é de benefativo. Em (ii) a construção é de verbo *faciendi*. João ajudou o irmão pagando o aluguel da casa dela. Nesse caso, não é relevante quem recebeu o aluguel.

(i) João pagou o aluguel ao locador.

(ii) João pagou o aluguel da casa para a irmã.

[19] O verbo é destacado em negrito, o OI é sublinhado e o OD vem em itálico.

2. leves

(38) **dá** *prejuízo* pa gente... (CZ-11)

(39) **dava** *um agrado* o menino (RC-09)

3. discendi

(40) eu vô **falá** *uma coisa* pa senhora (SP-08)

(41) eu vou **dizer** o senhô *que nem Tereza mesmo tá capaz de informar* (HV-12).

4. faciendi

(42) **comprá** *remédio* pos menino (SP-06)

(43) **fez** *um barraco* pr'ele (RC-09)

5. existenciais

(44) Num **tem** *uma menininha miudinha* pra mim (CZ-08)

(45) graças a Deus num **falta** *nada* pra mim (HV-22)

6. de condução

(46) mulé de tabinha **levô** *água* lá pra mim (HV-22)

(47) eu **mandava** pra essa mãe (RC-11)

7. de transferência

(48) O cacau **ia ficá** todo pa ela (SP-11)

(49) **distribuí** pra todo mundo (SP-03)

No amplo espectro semântico que caracateriza o universo de observação inicial desta análise, os OIs desempenham tanto os papéis de BENEFICIÁRIO, quanto de META (*finalidade*). Entretanto, os limites mais precisos da estrutura de dativo se restringiriam às construções em que o OI expressa o papel temático de BENEFICIÁRIO, devendo-se excluir aquelas em que o OI expressa o papel temático de META. Mas nem sempre a distinção entre um caso e outro é muito nítida, tendo por base apenas a interpretação semântica (cf. seção 18.1 supra). Considerando-se a taxonomia adotada, o papel temático de BENEFICIÁRIO estaria mais restrito às construções com os verbos benefativos, *discendi* e leves. E as estruturas em que o OI assume mais um papel de META seriam aquelas com os verbos *faciendi* ('fazer alguma coisa para alguém'), com os verbos existenciais ('ter alguma coisa para alguém'), com os verbos de condução ('levar alguma coisa para alguém'), com os verbos de transferência ('deixar alguma coisa para alguém'). Assim, essas últimas deveriam ser excluídas numa análise estrita da alternância dativa, restringindo-se o universo de observação para os verbos benefativos, *discendi* e leves. Mas, para que essa decisão não fosse tomada apenas com base no critério da interpretação semântica, buscou-se um parâmetro formal para fazer essa distinção.

Formalmente, os papéis semânticos de BENEFICIÁRIO e META estão mais associados ao emprego das preposições *a* e *para*, respectivamente. Porém, se essa relação não se tem dado de forma discreta desde a formação das línguas românicas (cf. seção 18.1), o crescente avanço da preposição *para* sobre a área da preposição *a* no PB (GOMES, 2003) só agravou o problema. Mesmo assim, alguns parâmetros do uso das duas preposições podem ser invocados, para buscar distinções formais que sustentem as distinções semânticas propostas. Como foi visto na seção 18.1, o uso exclusivo da preposição *para* só ocorre nos casos em que o OI desempenha o papel de META. O cruzamento desse fator com a taxonomia adotada apresentou resultados significativos. No *corpus* analisado, os verbos *faciendi*, existenciais, de condução e de transferência subcategorizaram somente OIs que podiam ser regidos apenas pela preposição *para*, configurando o papel de META desses argumentos. Entre os verbos benefativos, *discendi* e leves, predominaram os OIs que podiam ser regidos pelas duas preposições no PB. Só se registraram construções com OIs que apenas aceitavam a preposição *para* com os verbos *passar* (*um remédio*), que equivale a *receitar* e *fazer* (*uma leitura*), exemplificados em (50) e (51), respectivamente. Trata-se de dois verbos leves, que, como já foi dito anteriormente, recobrem uma gama ampla de significações. Nesse caso, assemelham-se aos verbos *faciendi*, no sentido de 'fazer alguma coisa para alguém'.[20] Portanto, o critério do uso exclusivo da preposição *para* revelou-se um parâmetro formal válido para sustentar a classificação semântica adotada e para justificar a exclusão das ocorrências com os verbos *faciendi*, existenciais, de condução e de transferência, numa análise mais precisa da alternância dativa.

(50) o médico **passô** *um remédio* pra ela (RC-13)

(51) **faz** *uma letchura* pra mim (CZ-09)

No plano empírico, o uso efetivo da preposição também apontou na direção do raciocínio aqui adotado. O emprego da preposição *a* nas comunidades de fala estudadas é muito raro, assim como o é no PPB em geral e na linguagem culta coloquial. Em um total de 289 OIs do *corpus* regidos por preposição, só 13 foram introduzidos pela preposição *a*. Ratificando a análise anterior, essas ocorrências se distribuíram apenas entre os verbos benefativos, *discendi* e leves.

Finalmente, ainda no plano empírico, os resultados da variável dependente confirmaram a delimitação proposta para o fenômeno da alternância dativa, ainda que se possa alegar um raciocínio tautológico nessa passagem. Nas estruturas com os verbos *faciendi*, existenciais, de condução e de transferência, só ocorreram CDPs, registrando-se um número de ocorrências razoavelmente representativo

---

[20] Note-se também que não ocorre a transferência de posse material, outra característica própria das construções de dativo. Esse parâmetro também será sistematicamente considerado adiante.

em todos esses contextos.[21] Esses resultados autorizam a exclusão desses fatores, pois a análise quantitativa da variação linguística prevê a exclusão dos contextos em que a ocorrência de uma variante é consistentemente categórica (PINTZUK, 1988; SCHERRE; NARO, 2003; TAGLIAMONTE, 2006).

Assim, com base, tanto em critérios semânticos e análiticos, quanto nos resultados empíricos do processamento quantitativo dos dados, a análise variacionista da forma como a estrutura de dativo se atualiza na gramática das comunidades de fala estudadas, configurando o fenômeno da alternância dativa, só considerou as estruturas com os verbos benefativos, *discendi* e leves, numa definição mais precisa do fenômeno em foco, na qual se conjugam dialeticamente a reflexão teórica disponível e os achados empíricos. Nesse escopo da alternância de dativo, pode-se dizer que os verbos benefativos são aqueles que determinam as estruturas de dativo *stricto sensu*, definindo a porção mais nuclear e legítima do fenômeno. Os verbos *discendi* desencadeiam uma estrutura semelhante, mas com especificidades diferenciadoras como a possibilidade bastante geral do OD ser uma oração completiva, o que inibe bastante a ordem OD-OI nesse contexto. E os verbos leves, por recobrirem uma gama bem ampla de significados, compreendem, tanto itens lexicais que se assemelham bastante aos verbos benefativos, quanto itens bem distintos, como se viu acima. Além disso, a alta coesão do verbo com o OD, com o qual forma uma única lexia composta, impede a ocorrência da ordem OI-OD.

Portanto, a mensuração da inversão da ordem normal na COD teve de ser refeita, considerando-se essas especificidades estruturais. A rigor, a análise da inversão só pode fundamentar-se empiricamente nas ocorrências com os verbos benefativos, pois a inversão da COD é obrigatória com os verbos leves (todas as 27 ocorrências com verbos leves do *corpus*, reunindo as ocorrências de CDP e COD, são com a ordem OD-OI, podendo-se dizer que essa seria a ordem categórica nesse contexto) e é bastante inibida com os verbos *discendi* por conta do princípio estrutural que desloca os constituintes pesados (no caso, os ODs formados por orações completivas) para o final da oração. Em todas as ocorrências do *corpus* em que o OD é uma oração, a ordem encontrada foi OI-OD, o que enviesa o resultado da inversão, tanto na CDP (que passa de 15% para 75%, nesse contexto), quanto na COD (que cai de 41% para 10%). Retirando, então, as ocorrências dos contextos que enviesam os resultados da inversão na COD, ao considerar esse fenômeno apenas junto aos verbos benefativos, chegou-se a uma frequência de 32% do total. A frequência caiu um pouco, mas ainda é significativamente representativa, o que mantém em pauta uma possível transferência gramatical a partir do substrato ewe.

---

[21] Foram 74 ocorrências com verbos *faciendi*, 26 de condução, 13 de transferência e 9 existenciais.

Dentro do universo em que ocorreu a **alternância dativa** na amostra de fala analisada, a variável tipo de verbo apresentou os resultados que figuram na Tabela 1.

Tabela 1 - Alternância dativa no português afro-brasileiro segundo a variável tipo de verbo

TIPO DE VERBO	CDP			COD		
	Nº. de oc. /Total	Freq.	P.R.	Nº. de oc. / Total	Freq.	P.R.
Benefativos	73/108	68%	**.40**	35/108	32%	**.60**
Leves	19/27	70%	**.45**	8/27	30%	**.55**
Discendi	80/99	81%	**.62**	19/99	19%	**.38**
Total	172/234	74%	—	62/234	26%	—

Nível de significância .045.

As orações com os verbos benefativos são o contexto em que a frequência da COD é mais elevada, atingindo 32% contra 26% de frequência geral (com peso relativo de .60). Os verbos leves aparecem em um nível intermediário, com uma frequência de COD de 30% e peso relativo de .55, bem próximo à neutralidade. Já os verbos *discendi* se revelaram o contexto mais refratário à alternância dativa, com uma frequência de apenas 19% de COD e peso relativo de .38.

Foi feita também uma nova rodada considerando a variável **transferência de posse material**. Na discussão apresentada na seção 18.1, esse fator foi tomado como um delimitador da alternância dativa. Contudo, esse critério não foi adotado aqui na delimitação do fenômeno, porque implicaria excluir os verbos *discendi*, em que não há transferência de posse material, mas a alternância dativa ocorre normalmente, inclusive no inglês (Ex.: *I've told Mary the truth*.). De qualquer forma, é válido considerar, em uma análise variacionista da alternância dativa, esse fator como favorecedor do fenômeno. Assim, considerando apenas os verbos benefativos e leves, os resultados dessa variável foram os constantes na Tabela 2.

Tabela 2 - Alternância dativa no português afro-brasileiro segundo a variável relação expressa pelo verbo

RELAÇÃO EXPRESSA PELO VERBO	CDP			COD		
	Nº. de oc./ Total	Freq.	P.R.	Nº. de oc./ Total	Freq.	P.R.
Transferência de posse	50/78	64%	**.44**	28/78	36%	**.54**
Outra	45/60	75%	**.57**	15/60	25%	**.43**
Total	172/234	74%	—	62/234	26%	—

Nível de significância .009.

Como se pode ver, a frequência de uso da COD sobe de 26% para 36% quando o verbo expressa uma relação de transferência de posse material, e os

pesos relativos confirmam essa relação como um fator que favorece a alternância dativa.[22]

A alta frequência da alternância dativa junto aos verbos benefativos e quando o verbo expressa uma transferência de posse material fornecem evidências empíricas importantes para a definição do escopo do fenômeno da alternância dativa. Por outro lado, a produtividade da COD nesses contextos, correspondendo a aproximadamente um terço das ocorrências, coloca o português afro-brasileiro como uma variedade linguística que exibe de forma consistente o fenômeno da alternância dativa. Ao apresentar um processo de reestruturação da gramática *sui generis* no quadro das tendências predominantes entre as línguas românicas, em geral, e no português, em particular, o português afro-brasileiro aproxima-se historicamente das línguas crioulas de base portuguesa da África, nas quais ocorreu um processo análogo.

Assim, considerando que a emergência da alternância dativa nessa variedade do PB está historicamente relacionada com o processo de transmissão linguística irregular desencadeado a partir da aquisição imperfeita do português como segunda língua por parte dos falantes adultos trazidos da África como escravos e da socialização dessa variedade de segunda língua nesse segmento, bem como a partir da nativização desse modelo defectivo entre os afrodescendentes, resta saber se o quadro do encaixamento social da alternância dativa nas comunidades de fala analisadas reflete essa perspectiva histórica.

18.3.2 O encaixamento social da alternância dativa nas comunidades rurais afro-brasileiras isoladas

Do ponto de vista da história sociolinguística das comunidades de fala analisadas, a emergência da alternância dativa teria ocorrido na sua fase de formação, e teria se conservado em uso durante um bom tempo, enquanto se manteve o seu isolamento, até que, ao longo do século XX, o seu uso entraria em declínio em função da crescente influência de padrões linguísticos adventícios (cf. capítulo 1 deste livro). A distribuição geracional do fenômeno nas comunidades de fala analisadas pode fornecer uma confirmação empírica significativa para essa hipótese, desde que os falantes mais velhos apresentem as maiores frequências de realização da COD, com um decréscimo progressivo à medida que se passe para os falantes

[22] As duas variáveis foram consideradas em rodadas separadas do VARBRUL, porque, reunidas, a variável *transferência de posse material* excluía a variável *tipo de verbo*, em função da superposição que existe entre elas. Isso explica os níveis de significância diferentes dos resultados dessas duas variáveis.

mais jovens. Os resultados da variável **faixa etária** ratificam tal hipótese, como se pode ver na Tabela 3.[23]

Tabela 3 - Alternância dativa no português afro-brasileiro segundo a variável faixa etária

FAIXA ETÁRIA	CDP			COD		
	Nº. de oc. / Total	Freq.	P.R.	Nº. de oc. / Total	Freq.	P.R.
20 a 40 anos	71/84	85%	**.69**	13/84	15%	**.31**
41 a 60 anos	62/78	79%	**.51**	16/78	21%	**.49**
Acima de 60 anos	39/72	54%	**.27**	33/72	46%	**.73**
Total	172/234	74%	—	62/234	26%	—

Nível de significância .009.

Os falantes com mais de sessenta anos são aqueles que mais empregam a COD, com uma frequência de 46%, que é quase o dobro da frequência geral (26%), e peso relativo de .73. Os falantes entre 41 e 60 anos situam-se em um nível intermediário, mas com uma frequência de uso da COD de apenas 21%, bem abaixo da encontrada entre os falantes mais velhos. De qualquer forma, o peso relativo de .49 indica a neutralidade escalar desse segmento. Já os falantes mais jovens exibem a menor frequência de uso da COD, ficando um pouco abaixo da faixa 2, com quinze pontos percentuais. E mais uma vez o peso relativo de .31 não deixa margem a dúvida sobre a significância da diferença de comportamento entre as faixas etárias.

A linha descendente exibida na Figura 1, que toma como base os pesos relativos, indica claramente um processo de **mudança em curso** no sentido do declínio do emprego da COD nas comunidades analisadas, de acordo com os princípios teórico-metodológicos da análise da mudança linguística em **tempo aparente** (LABOV, 1981; LUCCHESI, 2001b). Esse declínio seria determinado pela crescente influência dos padrões linguísticos urbanos sobre todas as regiões do país, nomeadamente a partir da segunda metade do século XX (cf. capítulo 1 deste livro). Tal influência tem desencadeado um processo de nivelamento linguístico, no qual submergem as marcas da fala popular que resultam de processos pretéritos de variação e mudança desencadeados pelo contato linguístico.

[23] Os pesos relativos atribuídos aos valores desta variável foram idênticos nas duas rodadas do VARBRUL feitas com as variáveis *transferência de posse material* e *tipo de verbo*.

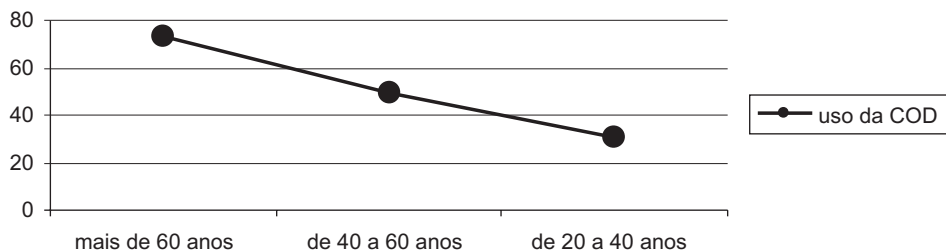



Figura 1 - Construções com objeto duplo segundo a faixa etária

A hipótese clássica do **tempo aparente** prevê que um falante consolida a sua gramática por volta dos quinze anos; dessarte, o comportamento linguístico de um falante de 60 anos refletiria a gramática de sua comunidade de fala há 45 anos (NARO, 2003). Considerando-se que, entre os falantes da faixa 3, há indivíduos com até cerca de cem anos, pode-se fazer uma projeção para a segunda década do século XX, quando as frequências de emprego da CDP e da COD seriam praticamente equivalentes. Isso significa que, no limiar do século XX, essas comunidades poderiam ser caracterizadas como variedades linguísticas de alternância dativa plena. Desde então, o uso da PDC tem avançado progressivamente, sobretudo a partir da década de 1950, de acordo com a retroprojeção da faixa 2, com base ainda na hipótese clássica. A tendência, então, é a de que a COD, que deriva da transmissão linguística irregular decorrente do contato linguístico que marcou a formação dessas comunidades, já tenha desaparecido na próxima geração de falantes dessas comunidades. Tal mudança, entretanto, implica o nascimento de indivíduos para os quais a COD seja uma construção agramatical. Completa-se, assim, o ciclo de uma mudança na gramática, em que alterações quantitativas no uso de uma forma linguística (no caso, a redução no uso da COD) atingem um ponto no qual a ocorrência dessa forma no *input* para aquisição da língua materna é tão baixo que tal estrutura não se integra na gramática das crianças que estão desenvolvendo sua língua materna. Nesse ponto, a mudança é **catastrófica**, no sentido de Lightfoot (1979, 1999), determinando uma clivagem na comunidade de fala entre os falantes que têm aquela forma (no caso, a COD) em seu repertório gramatical e aqueles novos falantes para os quais essa forma é agramatical (cf. capítulo 4 deste livro). Com isso, assume-se que a alternância dativa constituiria um valor de um parâmetro da estrutura de dativo da Gramática Universal (CHOMSKY, 1986). A presença da COD, ao lado da CDP, como opções para a estrutura de dativo na oração constituiria o valor positivo para a alternância dativa; enquanto a presença exclusiva da CDP como única forma disponível para atualizar a estrutura de dativo na oração constituiria o valor negativo.

A variável **comunidade de origem** também pode fornecer indicações sobre o desenvolvimento histórico do fenômeno bem como sobre sua distribuição diatópica. Os resultados dessa variável são apresentados na Tabela 4.

Tabela 4 - Alternância dativa no português afro-brasileiro segundo a variável comunidade de origem

COMUNIDADE	CDP			COD		
	Nº. de oc. / Total	Freq.	P.R.	Nº. de oc. / Total	Freq.	P.R.
Barra e Bananal	36/57	63%	**.30**	21/57	37%	**.70**
Helvécia	42/67	63%	**.32**	25/67	37%	**.68**
Cinzento	34/45	76%	**.57**	11/45	24%	**.43**
Sapé	60/65	92%	**.79**	5/65	8%	**.21**
Total	172/234	74%	—	62/234	26%	—

Nível de significância .009.

As comunidades de Barra e Bananal, situadas no município de Rio de Contas, e a comunidade de Helvécia, no Extremo Sul da Bahia, são aquelas em que a alternância dativa atinge um grau mais significativo, com 37% por cento de frequência da COD e peso relativo em torno de .70. A comunidade de Cinzento situa-se em um estágio intermediário, com 24% de emprego da COD e peso relativo de .43 (um pouco abaixo da neutralidade, situada em .50). Por fim, o processo de queda da COD estaria mais avançado em Sapé, com uma frequência de apenas 8% e peso relativo muito baixo (.21). A questão que se coloca, então, é saber se esses números refletem um nível diferenciado de reestruturação gramatical em cada comunidade ou a maior ou menor velocidade do processo de declínio da COD a partir do século XX.

A resposta a tal questão é difícil em face dos dados empíricos e registros históricos disponíveis. Há registros, por exemplo, de que a comunidade de Helvécia teria passado por um processo de reestruturação gramatical bem profundo e muito próximo da **crioulização** (cf. capítulo 2 deste livro). Isso poderia explicar o fato de Helvécia estar entre as duas comunidades que exibem um maior nível de alternância dativa. Não se dispõe de informações semelhantes sobre as demais comunidades. Portanto, não se pode propor uma explicação conclusiva para o fato de as comunidades de Rio de Contas, na Chapada Diamantina, apresentarem atualmente o mesmo grau de reestruturação gramatical. Essas comunidades talvez sejam as mais antigas do universo observado, pois a sua origem remonta ao início do século XVIII, sendo possível uma origem quilombola. Em compensação, talvez sejam as que mais tenham contato com o mundo exterior nas últimas décadas.

Por outro lado, considerando-se a distribuição geográfica das quatro comunidades, pode-se pensar que o avançado estágio de declínio da COD em Sapé

deve-se à sua maior proximidade da capital do Estado. Sapé situa-se no Município de Valença, um pouco ao sul do Recôncavo Baiano, cinturão agropecuário que envolve Salvador, mantendo uma estreita relação com essa cidade até meados do século XX. Desse modo, o processo de difusão dos modelos urbanos atingiria com mais intensidade as regiões mais próximas da capital do Estado, esgarçando-se à medida que avançasse para as regiões mais interioranas, o que poderia explicar a distribuição diatópica do fenômeno estudado.

Outra variável social cujos resultados quantitativos também apresentaram uma distribuição estatisticamente significativa foi a variável **escolaridade**.[24] Entre os falantes estudados, a distinção possível separa aqueles que tiveram algum convívio com a escolaridade (do que muitas vezes resulta apenas a habilidade de desenhar o próprio nome), definidos como **semianalfabetos**, daqueles que não tiveram qualquer contato com o sistema público de ensino, os **analfabetos**. Não obstante a precariedade desse sistema de educação pública, espera-se que aqueles que passaram por algum exercício de alfabetização sejam mais suscetíveis ao padrão linguístico dominante. Os resultados dessa variável confirmaram as expectativas, como se pode ver na Tabela 5.

Tabela 5 - Alternância dativa no português afro-brasileiro segundo a variável escolaridade

ESCOLARIDADE	CDP			COD		
	Nº. de oc. /Total	Freq.	P.R.	Nº. de oc. / Total	Freq.	P.R.
Analfabeto	63/100	63%	**.34**	37/100	37%	**.66**
Semianalfabeto	109/134	81%	**.62**	25/134	19%	**.38**
Total	172/234	74%	—	62/234	26%	—

Nível de significância .019.

A alternância dativa atinge um nível mais elevado entre os analfabetos, com uma frequência de COD de 37% (peso relativo de .66), e é mais baixa entre os semialfabetizados, com 19% de frequência da COD (peso relativo de .38). Pode-se perceber que, apesar de sua precariedade, o sistema de educação pública tem constituído um vetor de penetração dos padrões do comportamento linguístico hegemônico nas comunidades periféricas.

[24] Por um problema de distribuição dos informantes nas células das amostras, esta variável teve de ser rodada separadamente da variável *faixa etária*. O problema é que a maioria dos informantes semianalfabetos está na faixa 1 (de 20 a 40 anos), enquanto a maioria dos falantes mais velhos da faixa 3 são analfabetos. Essa superposição, que reflete a situação real da escolaridade nas comunidades estudadas, faz com que, no cruzamento da variável *faixa etária* com *escolaridade*, está última seja descartada. Retirando-se a variável *faixa etária*, a variável *escolaridade* é selecionada pelo VARBRUL como estatisticamente relevante. Isso explica o nível de significância diferente dos resultados desta variável.

O efeito da escolaridade deve ser compreendido no contexto mais amplo da relação com o mundo exterior, ou seja, a escolaridade é um meio de relacionar o indivíduo com a esfera do público, retirando-o do seu universo cultural doméstico. Portanto, o que está na base de tudo é o contato com um universo cultural mais amplo. Nesse sentido, seria de esperar também que os indivíduos que viveram mais de seis meses fora das comunidades apresentassem um menor grau de alternância dativa, porém os resultados dessa variável não tiveram significância estatística no processamento do programa das regras variáveis. O mesmo se deu com a variável sexo, cujos resultados também não exibiram uma distribuição estatisticamente relevante. Em relação a outras variáveis linguísticas analisadas nessas comunidades, os homens exibiram um comportamento mais próximo do padrão, em função de seu maior trânsito e maior contato com o mundo exterior; o que também se observa nas variedades populares na periferia das grandes cidades (BORTONI-RICARDO, 1985; RODRIGUES, 1992). Porém, os resultados do processamento quantitativo dos dados não ratificaram esse padrão em relação à alternância dativa.

Portanto, a hipótese de um processo de mudança em curso nas comunidades rurais afro-brasileiras que se caracteriza por um declínio da alternância dativa apoia-se basicamente nos resultados das variáveis sociais faixa etária e escolaridade.

18.4 Considerações finais

O fenômeno da alternância dativa, que se observa entre as línguas germânicas e está ausente no universo das línguas românicas, parece ser geral entre as línguas crioulas, mesmo entre aquelas que têm como língua lexificadora uma língua românica; o que oferece suporte empírico para correlacionar a emergência desse fenômeno à reestruturação gramatical que caracteriza os processos de transmissão linguística irregular desencadeado pelo contato entre línguas massivo, radical e abrupto. Apesar disso, não há evidências empíricas suficientes para determinar se a emergência da alternância dativa deriva-se de processos de transferência gramatical das línguas do substrato ou do processo de nativização da variedade linguística que se forma na situação de contato guiado pelas estruturas universais da faculdade da linguagem. A ampla distribuição do fenômeno em línguas crioulas de diferentes regiões do planeta, em princípio, entra em contradição com a hipótese do substrato africano. Por outro lado, não há evidências empíricas que comprovem que a COD seja adquirida independentemente do *input* disponível no processo de aquisição da língua materna, condição necessária para fazer da alternância dativa um valor não marcado na Gramática Universal.

De qualquer forma, a emergência da COD estaria certamente correlacionada a uma situação de pobreza de estímulo no que concerne à marcação morfológica do dativo. Nesse sentido, a fragilidade fonética da preposição *a*, que marca o dativo nas línguas românicas, constituiria um fator decisivo para a emergência da COD nas línguas crioulas delas derivadas. Por outro lado, o avanço da preposição *para* no PB, fazendo com que essa forma se tornasse uma preposição multifuncional, expressando tanto as relações de LOCATIVO e META, quanto a relação de BENEFICIÁRIO, teria inibido a emergência da COD nas variedades que se formaram com a transplantação da língua para o Brasil. A COD só se teria consolidado naquelas comunidades que se formaram em situações mais radicais de contato entre línguas, como é o caso das comunidades rurais afro-brasileiras isoladas. Nesse caso, o próprio emprego da preposição *para* como marcador de dativo seria defectivo o suficiente para dar ensejo à emergência da COD, havendo alguns indícios de uma possível transferência do substrato africano, o que abre uma perspectiva de desdobramento futuro dessa pesquisa, no sentido de definir de maneira mais precisa as formas como a estrutura de dativo se atualiza nas línguas mais representativas do substrato africano no Brasil, a saber: o quimbundo, umbundo e quicongo, na família banto; e o iourubá, ewe e fongbe, na família kwa.[25]

Portanto, a presença da COD no repertório sintático das comunidades rurais afro-brasileiras isoladas, bem como de outras variedades populares do PB ainda insuficientemente identificadas do ponto de vista sociolinguístico (SCHER, 1996), coloca a polarização sociolinguística do Brasil em um outro patamar, pois a COD é uma estrutura absolutamente estranha para os falantes da norma urbana culta do PB, sendo mesmo agramatical para a maioria deles. Dessa forma, pode-se dizer que a polarização, nesse caso, assenta em uma marcação paramétrica distinta. Assumindo uma marcação paramétrica para essa estrutura, a norma culta brasileira exibiria um valor negativo para a alternância dativa, enquanto, em algumas variedades do PPB, esse valor seria positivo, configurando-se a existência de duas gramáticas distintas, na norma culta e na norma popular. Confirma-se, assim também, a visão de que, se no plano socioeconômico a polarização sociolinguística do Brasil é o resultado de séculos de concentração de renda e da marginalização de largos segmentos da sociedade brasileira, essa polarização, no plano etnolinguístico, deriva de como o contato entre línguas teria afetado as diversas variedades do PB.

Nesse sentido, pode-se pensar que, nas variedades em que o contato entre línguas foi mais intenso e direto, como no caso do português afro-brasileiro aqui em foco, a erosão da morfologia de dativo foi tanta que permitiu a emergência da

[25] Uma outra forma de se perseguir essa hipótese do substrato seria tentar mapear diatopicamente o fenômeno da alternância dativa no Brasil, para estabelecer correlações a partir da presença do fenômeno em regiões com uma presença representativa desses grupos etnolinguísticos no passado.

COD, o que não ocorreu nas variedades cultas do PB, em que o incremento funcional da preposição *para* garantiu a marcação morfológica do dativo em níveis que impediram a emergência da COD. Contudo, não se pode descartar o efeito do contato entre línguas, mesmo quando se trata das mudanças ocorridas na norma culta do PB, como bem perceberam Morais, Ribeiro e Ferreira (no prelo).

O contato entre línguas afetou o comportamento linguístico dos falantes das classes da parte de cima da pirâmide social brasileira de forma indireta pelo que se pode chamar de **contato dialetal**. Ou seja, o comportamento dos falantes da norma culta brasileira não deriva de um processo de transmissão linguística irregular ocorrido no passado, mas do convívio com falantes de variedades derivadas desse processo de transmissão linguística irregular, ao longo de vários séculos, até os dias de hoje. Nesse sentido, pode-se pensar no convívio das crianças da classe dominante com amas de leite e escravos domésticos, bem como com os filhos de escravos, que costumavam brincar com os filhos dos senhores e dos capatazes, até serem incorporados na produção, a partir dos sete anos (cf. capítulo 1 deste livro). No presente, esse contato se reproduz com babás e empregadas domésticas provenientes do interior do país em função do êxodo rural.

Além disso, em um plano mais amplo da história, não se pode deixar de ter em conta o papel desempenhado por cerca de três milhões de imigrantes europeus e asiáticos que ingressaram na base da pirâmide social brasileira entre as últimas décadas do século XIX e primeiras décadas do século XX, adquirindo o português a partir dos modelos da fala popular de escravos e ex-escravos africanos crioulos e mestiços. Como esses imigrantes e seus descendentes na maioria dos casos ascenderam rapidamente na estrutura social brasileira, pode-se ver aí um forte vetor de circulação de estruturas da fala popular para a norma culta (LUCCHESI, 2001a e o capítulo 1 deste livro).

O enfraquecimento da preposição *a*, bem como do clítico de dativo *lhe*, pode ser tributado a esse efeito indireto do contato entre línguas na formação da norma culta brasileira, o que não deixa de ter paralelos com processos de crioulização, pois a emergência de uma preposição multifuncional, como *para* no PB, é uma das características das línguas crioulas que exibem a CDP. Entretanto, o efeito do contato entre línguas nas variedades cultas e semicultas do PB esgotou-se na erosão da morfologia privativa de dativo, não sendo suficiente para dar ensejo a um processo de reestruturação da gramática, como aconteceu no português afro-brasileiro com a emergência da COD.

Assim, pode-se retomar a distinção proposta por Lucchesi (2003), que identifica como sendo três os processos desencadeados pela transmissão linguística irregular: (i) *a erosão da morfologia flexional*, (ii) *a gramaticalização que gera novos mecanismos sintáticos* e (iii) *a mudança na marcação do valor do parâmetro na gramática*. No caso da norma culta, o único processo que a atingiu foi a erosão da morfologia

flexional. Já no caso do português afro-brasileiro, esta análise demonstrou que não se pode pensar nos três processos de forma estanque e discreta, e sim em sua interação dialética, pois, na emergência da alternância dativa no português afro-brasileiro, a erosão morfológica deu ensejo a uma reestruturação da gramática, do que resultou uma alteração no valor do parâmetro da estrutura de dativo.

Com isso, esta análise chega ao seu final com uma conclusão inexorável. A alternância dativa no português afro-brasileiro comprova a polarização socio-linguística do Brasil em um nível de radicalidade até então não conhecido, na medida em que implica uma diferença paramétrica que distingue a gramática da norma culta da gramática de algumas variedades do PB. Além disso, a análise desse fenômeno reuniu evidências irrefutáveis de que essa polarização socio-linguística deriva dos graus diferenciados com que o contato entre línguas afetou as diferentes variedades da língua portuguesa no Brasil.

19 A representação da primeira pessoa do plural

Dante Lucchesi

Este capítulo trata da variação na representação da 1ª pessoa do plural na função de sujeito, como exemplificado abaixo:

(1) *Nós* tava dentro do hotel! (RC-04)[1]

(2) *A gente* leva pa casa de farinha. (SP-04)

Em todas as variedades da língua portuguesa na América e na África, o pronome pessoal canônico da 1ª pessoa do plural *nós* enfrenta uma concorrência crescente da forma *a gente*, que constituía, antes de seu processo de **gramaticalização**, um Sintagma Nominal, formado pelo artigo e pelo nome, cujo significado era 'comunidade, população' (e.g., *a gente deste lugar é muito desconfiada*). Hoje, como resultado do processo de gramaticalização, a consciência dessa estrutura sintagmática original já se perdeu para a maioria dos brasileiros, sendo comum entre os estudantes do ensino fundamental e médio o "erro" de se grafar a expressão como uma única lexia: *agente*. Assim sendo, a forma *a gente* concorre com o pronome canônico *nós* e suas formas do caso oblíquo, em todas as funções sintáticas e em todos os níveis de referencialidade com que esse pronome é usado, desde a indeterminação do sujeito até o chamado *plural de modéstia*, em que o pronome *nós* refere-se exclusivamente à pessoa do falante.

No que concerne a essa variável linguística, o português afro-brasileiro exibe uma frequência de uso da forma inovadora *a gente* superior à que se observa na norma urbana culta, reforçando a ideia de que essa mudança tem se implementado de baixo para cima no português brasileiro (doravante PB), sendo liderada pelos segmentos populares de mais baixa escolaridade.

A análise variacionista desse fenômeno no universo das comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia é feita com base nos pressupostos teóricos do encaixamento linguístico e social dos processos de variação e mudança (WENREICH; LABOV; HERZOG, 1968 [2006]; LABOV, 1972a,

[1] Os exemplos extraídos do *corpus* analisado são indicados pelo número do informante e a sigla da comunidade a que esse informante pertence: CZ, Cinzento; HV, Helvécia: RC, Rio de Contas; SP, Sapé. Uma apresentação do *corpus* utilizado nesta análise será feita adiante.

1982, 1994). A sua base de dados foi retirada de uma amostra de fala constituída por 24 entrevistas de tipo sociolinguístico realizadas com moradores de quatro comunidades rurais afro-brasileiras isoladas de quatro diferentes regiões do Estado da Bahia, no período de 1992 a 2002. Todos os informantes tinham pouca ou nenhuma escolaridade, e alguns viveram por mais de seis meses fora da comunidade, trabalhando em centros urbanos. Na composição da amostra, os informantes foram distribuídos equitativamente por sexo e faixa etária. Desse modo, a amostra de fala foi estruturada de acordo com as seguintes variáveis: (i) comunidade: Sapé (Município de Valença, no Recôncavo Baiano), Helvécia (Município de Nova Viçosa, no Extremo Sul), Barra e Bananal (Município de Rio de Contas, na Chapada Diamantina), Cinzento (Município de Planalto, na região do Semiárido); (ii) sexo: masculino e feminino; (iii) idade: faixa 1, 20 a 40 anos; faixa 2, 41 a 60 anos; faixa 3, mais de 60 anos; (iv) escolaridade: semianalfabeto e analfabeto; (v) estada fora da comunidade: para figurar com valor positivo, o falante deveria ter vivido pelo menos seis meses fora da comunidade.[2]

Este capítulo está estruturado da seguinte maneira: inicialmente é traçado um breve panorama do processo diacrônico de gramaticalização de *a gente* e do quadro atual de variação nas diversas variedades do português brasileiro; em seguida, é analisado o encaixamento da variável na estrutura linguística das comunidades de fala analisadas; o encaixamento social é o objeto da análise da seção seguinte, a que se segue a conclusão.

19.1 A gramaticalização de *a gente* e o quadro atual da variação no Brasil

Em um estudo em tempo real de longa duração, Lopes (1999) identifica o início do processo de gramaticalização da expressão nominal *a gente* como pronome pessoal na língua portuguesa, no século XVI, mas só encontra ocorrências inequívocas para o uso de *a gente* como pronome pessoal nos séculos XVIII e XIX. No século XX, o processo de gramaticalização de *a gente* se consolida principalmente no português brasileiro e no português moçambicano, no continente africano.

A gramaticalização de *a gente* pode-se ter desencadeado a partir da lacuna deixada pelo desuso do substantivo *homem*, que, no português arcaico, funcionava como um pronome que indicava a indeterminação do sujeito, como o *on* do francês (também derivado do latim *hominem*). A perda dessa partícula gramatical no

[2] As comunidades analisadas são apresentadas de forma mais detalhada no capítulo 2 deste livro; e a metodologia empregada na constituição das amostras de fala é descrita no capítulo 5.

português teria ocorrido no século XVI. Assim, a gramaticalização da expressão nominal *a gente*, que em seu sentido original se referia a uma coletividade, pode ter se insinuado como forma de expressar um sujeito indeterminado, mas não há evidências empíricas que comprovem esse percurso — o que responderia ao que Weinreich, Labov e Herzog (1968 [2006]) definiram com o *problema da transição* —, pois, como já se disse aqui, Lopes (1999) só encontra ocorrências inequívocas do uso de *a gente* como pronome em textos dos séculos XVIII e XIX. Nessa altura, a expressão nominal *a gente*, gramaticalizada, concorre com o pronome canônico *nós*, mesmo no sentido mais específico deste; ou seja, o de *eu ampliado*: eu, o interlocutor e/ou outrem.

No Brasil, pode-se dizer que *a gente* já superou *nós* em frequência de uso na língua falada, em todas as variedades do PB, apesar de o pronome *nós* ainda predominar largamente na escrita, sobretudo nos textos formais. Na fala, mesmo na norma urbana culta, *a gente* predomina.

Na amostra linguística do português afro-brasileiro aqui analisada, a frequência geral de uso de *a gente* como pronome de 1ª pessoa do plural atinge praticamente três quartos do total de ocorrências, como se pode ver na Tabela 1.

Tabela 1 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro

PRONOME	Nº DE OCOR./TOTAL	FREQUÊNCIA
a gente	1328/1820	73%
nós	492/1820	27%

Esse amplo predomínio de *a gente* no português afro-brasileiro supera o que Lopes (1999, p.165) observou em uma amostra do Projeto da Norma Urbana Culta (Projeto NURC) constituída no início da década de 1970, pois, nessa amostra, *a gente* corresponde a apenas 42% do total de ocorrências. Numa amostra de fala com os mesmos informantes do Projeto NURC colhida na década de 1990, a frequência de uso de *a gente* sobe para 54%, contra 46% de *nós*, “evidenciando que o uso do *a gente* se torna mais frequente de uma década para outra”. Já o estudo de Omena (1996), realizado com o *corpus* do PEUL, que reúne entrevistas do tipo sociolinguístico com 48 falantes adultos do Rio de Janeiro com um a doze anos de escolaridade, encontrou um predomínio de *a gente* na ordem de 65% do total de ocorrências. Com falantes com as mesmas características da cidade de Florianópolis, Seara (2000, p.181) obteve um percentual de 70% de uso de *a gente*. Configuram-se, portanto, cenários bem semelhantes ao encontrado nas comunidades rurais afro-brasileiras.

Na análise da variação *nós* e *a gente* no português afro-brasileiro, cujos resultados são apresentados a seguir, buscou-se desvendar o encaixamento linguístico e social de um processo que também se afigura como uma mudança em

favor da forma inovadora, e que parece ocorrer em paralelo, com o que se tem observado na norma urbana culta e semiculta.

19.2 O encaixamento linguístico da implementação de *a gente* no português afro-brasileiro

No escrutínio do encaixamento linguístico da variação na forma do pronome da 1ª pessoa do plural na gramática das quatro comunidades rurais afro-brasileiras aqui analisadas, tendo por base o processamento quantitativo dos dados levado a cabo com o emprego do pacote de programas VARBRUL (PINTZUK, 1988; SCHERRE; NARO, 2003; TAGLIAMONTE, 2006), as variáveis linguísticas que se mostraram relevantes em termos estatísticos foram as seguintes: (i) nível de referencialidade; (ii) paralelismo discursivo; e (iii) tipo de texto. Os resultados de cada uma delas é apresentado nas subseções abaixo.

19.2.1 O nível de referencialidade

O uso do pronome de 1ª pessoa do plural no português abarca vários níveis de referencialidade, desde o seu significado básico, que compreende o falante, o(s) ouvinte(s) e/ou outrem, até o seu significado mais genérico, como índice de indeterminação do sujeito, passando pela possibilidade de referência ao próprio falante, exclusivamente, no que as gramáticas normativas designam *plural de modéstia* (e.g. CUNHA, 1981, p.205-206). Assim sendo, essa variável foi estruturada de acordo com os seguintes fatores:

(1) **eu + (você(s)) + (ele(s)) [+específico]**

Ex.: Aí *nós* foi tirá foto, *nós* tava em quatro, mas só tirô *nós* três. (CZ-01)

(2) **eu [+/- específico]**

Ex.: E *a gente*, quando teve televisão, *a gente* assistia e no ôto dia *a gente* já saía preocupado com trabalho. (HV-04)

(3) **indeterminação circunscrita [–específico]**

Ex.: Quando é bom de chuva aqui, a vez *a gente* faz aquelas bêrada, a vez o povo roça, faz aquelas bêrada pra plantá fejão, enche de milho, inda dá. (CZ-09)

(4) **indeterminação universal [–específico]**

Ex: Num adianta *a gente* insisti num caminho que não tem saída. (HV-12)

Em (1) temos o significado básico do pronome, em que o referente é composto por pessoas totalmente especificadas, em um espectro que inclui sempre o falante, ora acompanhado pelo(s) ouvinte(s), ora acompanhado por outra(s) pessoa(s), ora por ambos. O falante também pode usar *nós* ou *a gente* para se referir apenas a si mesmo, o que é exemplificado em (2). Nesse caso, a fronteira entre uma referência específica e uma referência mais genérica, abrangendo o grupo a que o falante pertence é, muitas vezes, fugidia.[3] No caso da referência genérica, definida aqui pelo traço semântico [–específico], foram distinguidos dois planos: de um lado, a referência genérica circunscrita ao grupo/comunidade do falante — cf. exemplo (3); de outro lado, a referência totalmente genérica, a que denominamos **indeterminação universal**, exemplificada em (4), em que o escopo da referência contempla qualquer ser humano. Como o entrevistador não faz parte da comunidade do falante, um parâmetro muito útil para distinguir esses dois últimos níveis de referência é verificar se esta inclui o entrevistador ou não. Os resultados de cada fator definido para essa variável constam da Tabela 2.

Tabela 2 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro segundo seu nível de referencialidade

REFERÊNCIA	*A GENTE*			*NÓS*		
	Nº de oc./Total	Freq.	P.R.	Nº de oc./Total	Freq.	P.R.
Indeterminação universal	85/99	86%	**.69**	14/99	14%	**.31**
Indeterminação circunscrita	396/495	80%	**.66**	99/495	20%	**.34**
O próprio falante	349/430	81%	**.59**	81/430	19%	**.41**
Grupo específico com falante	498/796	63%	**.33**	298/796	37%	**.67**
Total	1328/1820	73%	—	492/1820	23%	—

Nível de significância: .009.

Como se pode ver, *a gente* prevalece nos contextos de maior indeterminação, enquanto o contexto em que *nós* é mais frequente é exatamente o contexto de referência específica, com 37% de ocorrências e peso relativo de .67. Isso se deve ao fato já mencionado acima de o significado original da expressão *a gente* remeter a uma coletividade. Portanto, no espectro semântico da forma *a gente*, o traço [–específico] ainda se mantém forte. Essa correlação entre o uso de *a gente* e a referência mais genérica é geral, sendo também detectada nas análises variacionistas de outras variedades do PB (cf., por exemplo, OMENA, 1996, p.204; e LOPES, 1998, p.416).

De qualquer modo, dado o avançado grau de implementação de *a gente* na gramática das comunidades analisadas, essa variante predomina, em termos das frequências absolutas, em todos os níveis de referencialidade.

[3] Cf. Omena (1996, p.202): "é difícil detectar se o falante está se referindo à 1ª pessoa do singular ou do plural".

19.2.2 O paralelismo discursivo

O **paralelismo formal** tem sido uma variável bastante usada nas análises variacionistas tanto do português quanto de outras línguas (POPLACK, 1980; LEFEBVRE, 1981; SCHERRE, 1988, entre tantos outros). O conceito busca dar conta, segundo Scherre e Naro (1993, p.2), de "uma tendência de formas gramaticais semelhantes ocorrerem juntas (cf. SCHIFFRIN, 1981)". Desse modo, postula-se uma tendência do falante a repetir sua escolha ao longo de sua fala, tanto no interior da oração, no chamado **paralelismo formal**, quanto numa sequência de orações, no chamado **paralelismo discursivo** (SCHERRE; NARO, 1993). No plano do paralelismo discursivo, a variável foi estruturada segundo os fatores exemplificados abaixo, com a expectativa de que o uso de *a gente* fosse favorecido pela presença dessa forma na oração anterior, o mesmo se aplicando ao uso de *nós*:

(1) **primeira referência**: quando não havia uma menção anterior ao referente da forma pronominal empregada

(2) **referência anterior feita por *a gente* ou forma verbal não marcada**

Ex.: **A gente** num joga valeno nada não, *a gente* joga assim só po esporte mesmo. (SP-04)

Todo dia **a gente** ia. Mas num... num precisô *nós* ficá lá com ela, não, dano assistênça pra ela, não! (RC-13)

(3) **referência anterior feita por *nós* ou forma verbal marcada**

Ex.: **Nós** foi sábo, né, quato hora da tarde *nós* saiu daqui. (CZ-01)

Às vez, aquilo que **nós** achava que ia sê mais importante, assim, p*a gente* ir pa conhecê assim uns cantô, né? (CZ-01)

Os resultados da variável são apresentados na Tabela 3.

Tabela 3 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro segundo o paralelismo discursivo

FORMA DO PRONOME NA ORAÇÃO ANTERIOR	*A GENTE* Nº de oc./Total	Freq.	P.R.	*NÓS* Nº de oc./Total	Freq.	P.R.
Primeira referência	748/1028	73%	.47	280/1028	27%	.53
a gente ou sujeito não realizado com forma verbal não marcada	**550/616**	**89%**	**.70**	66/616	11%	.30
nós ou sujeito não realizado com forma verbal marcada	30/176	17%	.10	**146/176**	**83%**	**.90**
Total	1328/1820	73%	—	492/1820	23%	—

Nível de significância: .009.

O paralelismo discursivo mais uma vez revelou-se um poderoso fator condicionador das escolhas do falante. A presença de *a gente* ou de uma forma

verbal não marcada na oração anterior faz com que a frequência de uso dessa variante se eleve de 73% para 89% (com peso relativo de .70). No que concerne à forma *nós*, o condicionamento é ainda mais forte, fazendo com que o uso de *nós* supere o de *a gente*, quando na oração anterior o falante empregue esse pronome ou a forma verbal marcada com o morfema da 1ª pessoa do plural *-mos*. Nesse caso, o uso de *nós* passa de 23% para 83% (com peso relativo de .90). Essa correlação mais forte pode ser explicada pelo fato de *nós* ser a forma mais marcada, atuando mais fortemente no plano do paralelismo discursivo.

19.2.3 O tipo de texto

A variável tipo de texto foi estruturada contemplando o fato de o falante estar narrando um fato, descrevendo um objeto ou uma atividade, ou argumentando na defesa de um ponto de vista, com o intuito de detectar alguma correlação entre o tipo de texto e a escolha do pronome da 1ª pessoa do plural. A variável foi, então, constituída pelos seguintes fatores:

(1) **Texto argumentativo**

Ex.: tudo que Deus dá, *a gente* tem que recebê aquela... aquela graça que Deus dá. (CZ-12)

(2) **Texto descritivo**

Ex.: *A gente* leva pa casa de farinha, raspa a mandioca e... é uma mã de obra, mas *a gente* faz assim mesmo, tem um forno, *a gente* usa lá, de lenha... Aí, só é jogá o fogo, e jogô a massa, liga o motô e dêxa rolá inté ti... a... a... a farinha tivé pronta. (SP-01)

(3) **Texto narrativo**

Ex.: Cheguei, troquei, tirei aquela roupa, que era muita roupa, dei ela de mamá, aí depois *nós* saiu. *Nós* fomo pa casa da vizinha, aí cheguei na hora... Aí ela... ela... deu na hora de *nós* 'panhá ônibus, aí *nós* veio embora. (HV-07)

Os resultados figuram na Tabela 4.

Tabela 4 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro segundo o tipo de texto

TIPO DE TEXTO	*A GENTE*			*NÓS*		
	Nº de oc./Total	Freq.	P.R.	Nº de oc./Total	Freq.	P.R.
Argumentativo	99/116	85%	**.64**	17/116	15%	**.36**
Descritivo	820/999	82%	**.54**	179/999	18%	**.46**
Narrativo	403/694	58%	**.42**	291/694	42%	**.58**
Total	1328/1820	73%	—	492/1820	23%	—

Nível de significância: .009.

A predominância de *a gente* nos textos argumentativos e descritivos está correlacionada ao traço semântico [–específico] dessa variante, pois, nesses tipos de texto, o falante refere-se, normalmente, a indivíduos em geral, seja na formulação de situações hipotéticas que fundamentem a sua argumentação, seja na descrição genérica de processos. Já na narrativa, que tem por objeto fatos particulares com agentes específicos, o uso de *nós* aumenta.

19.3 O encaixamento social da implementação de *a gente* nas comunidades de fala analisadas

No plano do encaixamento social, foram propostas as seguintes variáveis: (i) faixa etária; (ii) sexo; (iii) escolarização; (iv) estada fora da comunidade; e (v) comunidade (conforme apresentado acima). Dessas, o programa das regras variáveis, VARBRUL, selecionou as variáveis faixa etária, comunidade e escolarização como estatisticamente relevantes. Os resultados e a análise de cada uma dessas variáveis são apresentados a seguir.

19.3.1 A faixa etária

A variável faixa etária é crucial na análise sociolinguística dos processos de variação e mudança no que se convencionou chamar de **tempo aparente**, pois as diferenças entre os falantes de diferentes gerações são tomadas como o reflexo das diferenças na gramática da comunidade ao longo do tempo (LABOV, 1981; LUCCHESI, 2001b, entre outros). Tendo distribuído os informantes da amostra em três faixas etárias, que correspondem a três gerações distintas, foram encontrados os resultados que estão na Tabela 5.

Tabela 5 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro segundo a faixa etária do falante

FAIXA ETÁRIA	*A GENTE*			*NÓS*		
	Nº de oc./Total	Freq.	P.R.	Nº de oc./Total	Freq.	P.R.
Faixa 1: 20 a 40 anos	712/818	87%	**.74**	106/818	13%	**.26**
Faixa 2: 41 a 60 anos	393/590	67%	**.35**	197/590	33%	**.65**
Faixa 3: + de 60 anos	223/412	54%	**.24**	189/412	46%	**.76**
Total	1328/1820	73%	—	492/1820	23%	—

Nível de significância: .009.

Os resultados quantitativos revelam que o uso de *a gente* é maior entre os falantes mais jovens, decaindo na medida em que se passa para as faixas dos falantes

de idade mais avançada. O cenário de uma curva ascendente, como mostrado na Figura 1, aponta para um processo de mudança em curso, de substituição da variante canônica *nós* pela variante inovadora *a gente* nas comunidades rurais afro-brasileiras isoladas.

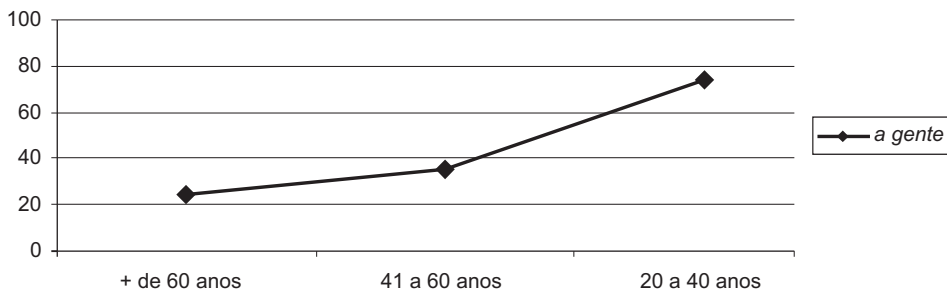



Figura 1 - Uso de *a gente* como pronome de primeira pessoa do plural no português afro-brasileiro segundo a faixa etária do falante (em peso relativo)

O cenário aqui apresentado para as comunidades rurais afro-brasileiras isoladas se ajusta ao que se tem observado em diversas variedades do PB, indicando fortemente uma mudança em curso em favor da implementação da forma inovadora *a gente* na fala dos brasileiros em geral. Analisando a fala popular do interior do Estado do Rio de Janeiro, Machado (1997, p.20) afirma que "a estratificação etária obtida demonstra haver, nos dialetos norte-fluminenses, uma propensão geral à substituição do pronome *nós* por *a gente*". Em uma amostra de falantes com nível de escolaridade fundamental e médio da cidade do Rio de Janeiro recolhida no início da década de 1980, Omena (1996, p.312-313) observou que a frequência de uso de *a gente* cresce de 51%, na fala de indivíduos com mais de 50 anos, para 67%, na fala de indivíduos entre 26 e 49 anos, atingindo uma frequência de 87% na fala de indivíduos de 15 a 25 anos — frequência bem próxima da encontrada na fala de crianças de sete a quatorze anos, 90%.[4] Segundo a autora, "esses resultados indicam que os falantes nascidos a partir de aproximadamente 1960 usam bem mais a forma *a gente*". Em um estudo de falantes com esses mesmos graus de escolaridade da cidade de Florianópolis, Seara (2000) observou também um padrão crescente de uso da forma *a gente*, porém sem a mesma intensidade: 72% para os falantes com mais de 50 anos; 70% para os falantes de 25 a 50 anos; e 76% para os falantes de 15 a 24 anos.[5]

---

[4] Os pesos relativos confirmam a implementação da forma inovadora à proporção que se passa para as novas gerações: .22 para os falantes com mais de 50 anos; .36 para os falantes de 26 a 49 anos; .67 para os falantes de 15 a 25 anos; e .74 para os falantes de sete a quatorze anos.

[5] Os pesos relativos indicam mais claramente um padrão ascendente: .40 para os falantes com mais de 50 anos; .51 para os falantes de 25 a 50 anos; e .69 para os falantes de 15 a 24 anos.

Entretanto, em sua análise em tempo real de curta duração, com base em duas amostras do Projeto NURC, com falantes do Rio de Janeiro com curso superior completo, entrevistados na década de 1970 e de 1990, Lopes (1999, p.189), apesar de reconhecer que "a substituição de *nós* por *a gente* esteja sendo implementada de forma acelerada nos últimos vinte anos no português do Brasil", chega à conclusão de que o processo "se caracteriza dentro dos modelos interpretativos de Labov (1994) como um padrão de gradação etária, pois se configura, no *estudo de tendências*, um comportamento estável da comunidade, e no *estudo painel* um comportamento instável dos mesmos indivíduos". Em um outro estudo em tempo real de curta duração, Omena (2003, p.79-80) compara os resultados obtidos em sua amostra do início da década de 1980 com uma amostra com as mesmas características recolhida entre 1999 e 2000, e também não chega a um resultado conclusivo. Segundo a autora, "quando se compara o desempenho dos falantes, em grupo, nos diferentes momentos, podemos concluir que a variação *nós / a gente* apresenta a mesma proporção de ocorrência através do tempo, a forma inovadora continua a predominar e a direção da mudança é a mesma". Entretanto, afirma que "os resultados demonstram uma situação de estabilidade, no presente momento". A falta de clareza nos diagnósticos dessas abordagens em tempo real de curta duração pode decorrer das limitações metodológicas que esse tipo de estudo ainda enfrenta,[6] pois também não se apresentou nenhuma evidência clara de uma situação de *variação estável*.

Um índice de variação estável seria uma resistência dos falantes à variante inovadora no plano da avaliação subjetiva das variantes. Porém, os resultados da variável escolaridade parecem indicar que a resistência ao uso de *a gente* na fala é cada vez menor no Brasil.

19.3.2 A escolaridade

No estudo de Omena (1996, p.315-319), a relação entre o uso do pronome da 1ª pessoa do plural com a escolaridade apresenta uma dupla face. De um lado, observa-se uma significativa diferença entre as crianças do curso primário e do ginasial, já que as últimas usam muito mais a variante padrão do que as primeiras, revelando que a escola reforça o uso da forma conservadora, na primeira fase da

[6] Cf. Omena (2003, p.79-80): "o período de tempo decorrido entre as duas investigações, mais ou menos 20 anos, é irrisório, o que dificulta a compreensão do processo". Além disso, "a má distribuição das amostras, resultante dos problemas enfrentados no recontato, pode ter enviesado os resultados".

escolarização. Por outro lado, entre os falantes adultos, os indivíduos com mais alto grau de escolarização são aqueles que menos usam a forma *nós*. Seara (2000, p.190-191) encontra o mesmo quadro em Florianópolis, onde "à medida em que se aumenta o grau de escolaridade, se tende a um maior uso da variante *a gente*" quando "se esperava que, com o aumento da escolaridade, se tivesse justamente o contrário".

Nas comunidades rurais afro-brasileiras aqui analisadas, são igualmente os falantes com alguma escolaridade aqueles que mais fazem uso da variante inovadora *a gente*, como se pode ver na Tabela 6.

Tabela 6 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro segundo o nível de escolaridade do falante

NIVEL DE ESCOLARIDADE	*A GENTE*			*NÓS*		
	Nº de oc./Total	Freq.	P.R.	Nº de oc./Total	Freq.	P.R.
Semianalfabeto	727/907	80%	**.56**	180/907	20%	**.44**
Analfabeto	601/913	66%	**.44**	312/913	34%	**.56**
Total	1322/1809	73%	—	487/1809	23%	—

Nível de significância: .009.

Os falantes que têm qualquer contato com o universo da escola (muitos só sabem assinar o nome) usam menos a variante conservadora *nós* do que os completamente analfabetos. É claro que o papel da escolaridade nesse universo é bastante distinto do que se observa nos grandes centros urbanos. Nas comunidades rurais, a experiência com a escolaridade reflete uma maior relação com o mundo exterior, além do universo do trabalho braçal do campo. Portanto, é natural que os falantes com alguma escolarização liderem um processo de mudança em sintonia com o que está ocorrendo nos grandes centros urbanos.

Por outro lado, o processo de variação e mudança na forma do pronome da 1ª pessoa do plural parece apresentar uma distribuição espacial diferenciada no território brasileiro, conforme já havia sido observado por Lopes (1998, p.418). Essa variação diatópica pode refletir o processo de difusão espacial da mudança, como se pode ver na análise da variável comunidade.

19.3.3 As comunidades

Considerando as quatro comunidades rurais afro-brasileiras aqui analisadas, que se situam em diferentes regiões do Estado da Bahia, os resultados dessa variável indicam que a implementação da variante *a gente* nas comunidades rurais afro-brasileiras tem uma distribuição diatópica significativa, como se poderá ver a partir dos números da Tabela 7.

Tabela 7 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro por comunidade isolada

NIVEL DE ESCOLARIDADE	*A GENTE*			*NÓS*		
	Nº de oc./Total	Freq.	P.R.	Nº de oc./Total	Freq.	P.R.
Sapé	334/337	**99%**	.96	03/337	**01%**	.04
Helvécia	357/499	**72%**	.43	142/499	**28%**	.57
Barra e Bananal	280/357	**78%**	.36	77/357	**22%**	.64
Cinzento	357/627	**57%**	.23	270/627	**43%**	.77
Total	1328/1820	73%	—	492/1820	23%	—

Nível de significância: .009.

Como também se pode ver na Figura 2, o processo de substituição de *nós* pela forma *a gente* está praticamente concluído na comunidade de Sapé, estando bem avançado nas comunidades de Helvécia e Barra e Bananal; em Cinzento, a variante *nós* ainda guarda alguma vitalidade.

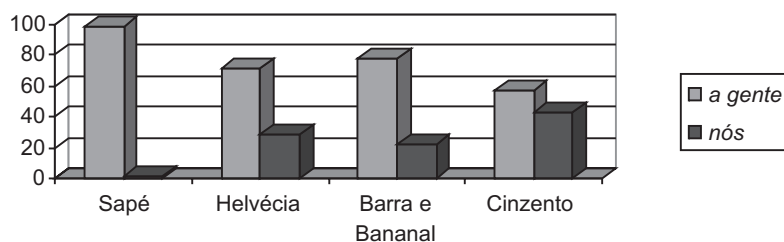



Figura 2 - Forma do pronome de primeira pessoa do plural no português afro-brasileiro por comunidade isolada (em percentuais)

Esses resultados reforçam a hipótese de que a implementação da forma *a gente* teria origem na região do Recôncavo Baiano, expandindo-se paralelamente pelo litoral, e pela rota da pecuária e da mineração para o interior do estado. Em seu estudo, Lopes (1999) analisou o processo de implementação de *a gente* em três variedades do português: a variedade europeia, o português brasileiro e o português de Moçambique, na África, e observou que o processo está mais adiantado nas duas últimas, o que aponta para a hipótese de que tal mudança seria favorecida pelas situações de contato entre línguas. Considerando-se que, até o século XIX, o Recôncavo concentrava a maior parte da população escrava da Bahia, é plausível que esse processo tenha ganho um grande impulso nessa região.

19.4 Conclusão

A análise quantitativa da variação na representação da 1ª pessoa do plural em quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da

Bahia focalizando o seu encaixamento linguístico não obteve resultados diferentes dos já alcançados na análise de outras variedades do PB. Apesar de predominar em todos os níveis de referencialidade, a variante inovadora *a gente* é ainda mais frequente nos contextos de referência genérica, o que está associado ao significado original dessa forma, que remete à ideia de 'coletividade'. Como consequência dessa correlação, *a gente* é a forma preferida em textos argumentativos e descritivos, quando o falante se refere genericamente a processos e sujeitos não especificados, enquanto o uso de *nós* aumenta em textos narrativos, que se referem a eventos e agentes específicos. O paralelismo discursivo também se revelou um fator condicionador da escolha do falante, pois, se ele seleciona a forma *a gente*, e sobretudo a forma *nós*, numa determinada oração, muito provavelmente fará a mesma escolha na oração seguinte.

No plano do encaixamento social, o quadro que emergiu foi o de uma mudança em curso em favor de *a gente*, com uma curva ascendente no gráfico da variável faixa etária, demonstrando que os mais jovens são os que mais usam a forma inovadora e os mais velhos aqueles que mais conservam o uso do pronome canônico *nós*. Essa mudança é liderada pelos indivíduos com algum contato com o universo da escolarização, revelando que não há uma avaliação negativa do uso de *a gente* nas comunidades analisadas.

Por outro lado, uma maior frequência de uso de *a gente vis-à-vis* às variedades urbanas do PB não indica uma mudança de cima para baixo, através de um processo de difusão linguística a partir dos grandes centros urbanos, como se observa em relação a outras variáveis linguísticas. Ao contrário, a implementação da forma *a gente* parece constituir uma mudança linguística de baixo para cima, sendo liderada pelas variedades populares do PB. O cotejo de frequência entre as comunidades analisadas aponta para um processo de difusão da mudança a partir do Recôncavo Baiano, zona que se situa em torno da capital e da Baía de Todos os Santos e concentra uma grande população de origem africana, em função dos engenhos de cana-de-açúcar e das plantações de fumo que ali vicejaram, entre os séculos XVII e XIX. A partir daí, a mudança se espalhou pelo litoral em direção ao Sul, e pelo interior seguindo a rota da pecuária e da mineração.

A origem da mudança na região do Recôncavo Baiano reforça a ideia de que o contato entre línguas teria impulsionado a substituição de *nós* por *a gente*, o que se coaduna com o fato de esse processo estar mais avançado no Brasil e em Moçambique do que em Portugal. Esse tipo de substituição não corresponde a uma mudança típica de situações de contato maciço e radical entre línguas, tanto que não ocorre em qualquer crioulo português da África. Portanto, se o contato entre línguas pode ter impulsionado a substituição do pronome *nós* pela forma *a gente*, certamente não foi o seu fator desencadeador.

20 A flexão de caso dos pronomes pessoais

Dante Lucchesi
Elisângela dos Passos Mendes

Dentre os aspectos da estrutura da gramática que diferenciam o português brasileiro do português europeu está a propriedade de os pronomes pessoais assumirem formas diferentes consoante a função sintática que desempenhem, a chamada **flexão de caso dos pronomes pessoais**, que se manteve na passagem do latim ao português, diferentemente do que ocorreu com os nomes. No Brasil, a substituição dos pronomes *tu* e *vós* da 2ª pessoa do discurso pela forma *você(s)* e do pronome da 1ª pessoa do plural *nós* pela forma *a gente* levaram a uma forte simplificação morfológica, já que essas formas, de origem nominal, não se flexionam quanto ao caso. Essa redução da flexão de caso dos pronomes pessoais é ainda acentuada no Brasil com o uso crescente da forma de sujeito do pronome da 3ª pessoa nas funções de objeto direto e indireto. Já em Portugal, a grande vitalidade de *tu* em todo o país (e de *vós*, em algumas regiões do Norte) e a baixa produtividade de *a gente* mantêm quase intacta a flexão de caso dos pronomes pessoais, até porque também não se observa em Portugal a substituição dos clíticos *o*(*a/os/as*) e *lhe*(*s*) pela forma do caso reto *ele*(*a*)(*s*).

Entretanto, esse cenário geral de simplificação da flexão de caso no português brasileiro comporta uma clivagem. Essa clivagem que, de resto perpassa toda a estrutura da língua no Brasil, distingue o processo histórico de formação dos padrões coletivos de comportamento linguístico dos segmentos do topo da sociedade brasileira, que têm acesso à escolaridade e ao universo institucional da cidadania (a chamada **norma culta**), do processo de formação das suas variedades populares (a **norma popular**), no que Lucchesi (2001, 2002a, 2002b, 2006a) denominou **polarização sociolinguística do Brasil**.[1] Desde os primeiros séculos da colonização, a elite do Brasil buscava manter a sua fala em conformidade com o padrão da metrópole portuguesa. Essa influência normatizadora de Portugal sobre as classes médias e altas da sociedade brasileira só arrefeceu a partir do século XX, apesar de as gramáticas normativas brasileiras ainda reproduzirem os modelos do português

[1] Veja-se também o capítulo 1 deste livro.

europeu. As variedades populares têm a sua origem nas situações de contato do português com as línguas indígenas e africanas. Tais situações produziram alterações significativas na estrutura da língua portuguesa, sobretudo através de processos de simplificação morfológica. As diferenças históricas na formação das duas grandes variedades do português brasileiro determinaram processos distintos de estruturação do sistema de flexão de caso dos pronomes pessoais.

Na norma culta, a redução na flexão de caso dos pronomes, como já se disse, está relacionada principalmente com a entrada das formas *você* e *a gente* na pauta dos pronomes pessoais, processo que ganha corpo provavelmente a partir da segunda metade do século XIX. A simplificação da flexão de caso nas variedades populares do português brasileiro remonta aos primeiros séculos da colonização do Brasil e decorre do processo de transmissão linguística irregular desencadeado pelo contato do português com as línguas indígenas e africanas (cf. capítulo 3 deste livro). Assim, esse processo se assemelha ao que ocorreu nas línguas crioulas de base lexical portuguesa da África, em que o sistema de flexão de caso dos pronomes pessoais da língua lexificadora foi praticamente eliminado. Para abordar essas questões, este capítulo é composto por uma descrição do sistema de flexão de caso dos pronomes pessoais na norma culta do português brasileiro, a que se seguem uma exposição sobre esse mesmo tópico em três línguas crioulas de base lexical portuguesa da África e uma descrição da variação na flexão de caso dos pronomes pessoais, a partir dos dados de uma amostra de fala colhida em quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia. Na conclusão, é feita uma síntese dos dados apresentados para demonstrar como a diferença no espectro da variação que separa a norma culta da norma popular é o reflexo de dois processos históricos distintos de variação e mudança que caracterizam as duas grandes vertentes do português brasileiro.

20.1 A flexão de caso dos pronomes pessoais na norma culta brasileira

Considerando-se apenas os falantes urbanos com alto grau de escolaridade do português brasileiro (doravante PB), os falantes da chamada **norma urbana culta**, já é possível observar uma instabilidade no que concerne à flexão de caso dos pronomes pessoais. Desse modo, a norma culta brasileira revela uma realidade bem diferente da contida no que se pode chamar de **norma padrão**.[2] Em geral, as gramáticas normativas apresentam a flexão de caso dos pronomes pessoais através

[2] Para uma distinção entre norma culta e norma padrão no Brasil, veja-se Lucchesi (2002b).

da distinção entre as formas do **caso reto**, que desempenham a função de sujeito, e as formas do **caso oblíquo**, que desempenham as funções de complementos verbais e nominais e de adjuntos adverbiais. Estas últimas podem ser **tônicas** e **átonas**; aquelas são regidas por preposição, e estas ligam-se diretamente ao verbo cujo sentido completam. Quando se trata da preposição *com*, são apresentadas formas contraídas. Esses pronomes estão reunidos no Quadro 1.

PESSOA DO DISCURSO	CASO			
	reto	oblíquo átono	oblíquo tônico	contrações
1ª pessoa do singular	*eu*	*me*	*mim*	*comigo*
2ª pessoa do singular	*tu*	*te*	*ti*	*contigo*
3ª pessoa do singular	*ele/ela*	*a/o, lhe, se*	*ele/ela, si*	*consigo*
1ª pessoa do plural	*nós*	*nos*	*nós*	*conosco*
2ª pessoa do plural	*vós*	*vos*	*vós*	*convosco*
3ª pessoa do plural	*eles/elas*	*os/as, lhes, se*	*eles/elas, si*	*consigo*

Quadro 1 - Flexão de caso dos pronomes na norma padrão do português

As formas oblíquas átonas desempenham normalmente as funções de objeto direto e indireto (e.g., *Maria **te** viu ontem*; *ele **nos** disse a verdade*). Para a 1ª e a 2ª pessoas, uma única forma desempenha as duas funções (e.g., *Maria **me** viu ontem*; *Maria **me** disse a verdade*), enquanto a 3ª pessoa tem uma forma para a função de objeto direto (**caso acusativo**) e outra para a função de objeto indireto (**caso dativo**), como se pode ver nos exemplos abaixo:

(1) A Maria estava na faculdade, eu **a** vi lá.

(2) A Maria me pediu, e eu **lhe** contei toda a história.

As formas oblíquas tônicas, sempre regidas por uma preposição, desempenham as funções de complemento oblíquo,[3] adjunto adverbial ou complemento nominal, como exemplificado em (3), (4) e (5), respectivamente.

(3) Ela não gosta **de vós**.

(4) a. Ela fez esse bolo **para nós**.

b. Ela não vai sair **contigo**.

(5) Quanto **a mim**, não há problema.

[3] A tradição gramatical classifica como **objeto indireto** todo complemento verbal regido por uma preposição. Porém, adotou-se, nesta análise, a distinção entre o **objeto indireto**, complemento verbal preposicionado que pode ser cliticizado (e.g., *dei o livro **à Maria** ~ dei-**lhe** o livro*), e o complemento oblíquo, o complemento verbal preposicionado que não pode ser cliticizado (e.g., *gosto **da Maria** // *gosto-**lhe***).

As formas oblíquas, tanto átonas quanto tônicas, da 1ª e da 2ª pessoas podem assumir um valor **reflexivo** ou **recíproco** (e.g., *tu **te** feriste*; *nós só dependemos **de nós***); já a 3ª pessoa tem formas exclusivas para expressar reflexividade e reciprocidade (e.g., *ele **se** feriu*; *eles só dependem **de si***). Portanto, mesmo no padrão normativo, só a 3ª pessoa mantém uma maior especialização de formas. Mesmo assim, ocorre o sincretismo nas funções de sujeito, complemento oblíquo e adjunto adverbial (e.g., ***ela** não veio*; *eu não gosto **dela***; *eu saí **com ela** ontem*, respectivamente).

Já o sistema de flexão de caso dos pronomes pessoais na norma culta brasileira, sobretudo em sua modalidade oral, é bastante diferente do que está prescrito nas gramáticas normativas, o que levou Monteiro (1991, p.118), que realizou uma descrição desse sistema com base nos materiais do Projeto NURC, a afirmar que ele se encontra "em fase de desestruturação". Grande parte das alterações e flutuações observadas na fala dos indivíduos escolarizados das grandes cidades decorre da substituição das formas *tu* e *vós* por *você(s)*, já consumada na grande maioria do território brasileiro, e da substituição de *nós* por *a gente*, bastante avançada em todas as normas do PB, sobretudo na modalidade falada (cf. capítulo anterior deste livro).

A introdução dessas formas nominais teve como consequência imediata uma forte redução no sistema de flexão de caso, na medida em que elas podem figurar em todas as posições sintáticas, sem se flexionar. Entretanto, essas formas coocorrem com as formas oblíquas, estabelecendo-se o quadro de variação exemplificado em (6) e (7):

(6) a. Você me viu, mas eu não vi **você** ~ mas eu não **te~lhe** vi.

b. Você me pediu e entreguei o documento **pra você** ~ e **lhe/te** entreguei o documento.

c. Você quer, mas eu não vou sair **com você** ~ **contigo**.

(7) a. A gente se apresentou, mas não escolheram **a gente** ~ mas não **nos** escolheram.

b. A gente viu a Maria, mas ela não viu **a gente** ~ mas ela não **nos** viu.

c. A gente pediu, mas a Maria não saiu **com a gente** ~ não saiu **conosco**.[4]

Além desses casos, deve-se registrar o virtual desaparecimento do clítico acusativo de 3ª pessoa na gramática natural dos brasileiros, sendo o seu uso o resultado da ação normativa da escola (cf. capítulo 17 deste livro). Na fala distensa, usa-se normalmente a forma do caso reto na função de objeto direto (doravante OD): *eu encontrei **ela** no cinema*. Na função de objeto indireto (doravante OI), a forma do caso reto também concorre com *lhe*, por conta da ambiguidade decorrente

---

[4] A forma *conosco* não é muito frequente no PB.

do emprego de *lhe* com referência à 2ª pessoa: *eu entreguei o livro* ***pra ela*** ~ *eu* ***lhe*** *entreguei o livro*.

Deve-se destacar, porém, a vitalidade das formas reflexivas e recíprocas (e.g., *ele* ***se*** *feriu*; *eles* ***se*** *desentenderam*), estendidas para a 2ª pessoa com a introdução de *você* (e.g., *você* ***se*** *feriu*; *vocês* ***se*** *desentenderam*).

O sistema de flexão da norma culta do PB está esquematizado no Quadro 2.

PESSOA DO DISCURSO	FUNÇÃO SINTÁTICA		
	sujeito	OD / OI	complemento oblíquo/ adjunto adverbial
1ª pessoa do singular	*eu*	*me*	*mim, comigo*
2ª pessoa do singular	*você*	*você ~ te ~ lhe / se*	*você, contigo*
3ª pessoa do singular	*ele/ela*	*ele/ela ~ o/a ~ lhe / se*	*ele/ela*
1ª pessoa do plural	*a gente ~ nós*	*a gente ~ nos*	*a gente ~ nós, conosco*
2ª pessoa do plural	*vocês*	*vocês ~ lhes / se*	*vocês*
3ª pessoa do plural	*eles/elas*	*eles/elas ~ os/as ~ lhes / se*	*eles/elas, si*

Quadro 2 - Flexão de caso dos pronomes pessoais na norma culta brasileira

Apesar do amplo processo de variação, pode-se constatar que a perda de flexão de caso na norma culta brasileira decorre, por um lado, da substituição de pronomes canônicos por expressões nominais, que já não possuíam flexão de caso (*você* e *a gente*), e, por outro lado, de fatores estruturais paradigmáticos, como se verá agora.

Em um texto clássico, Câmara Jr. (1972, p.47-53) não viu o uso corrente de uma forma invariável do pronome de 3ª pessoa em todas as funções sintáticas no português brasileiro como o reflexo de uma tendência evolutiva da língua já presente no português arcaico, pois, ao contrário do uso esporádico e enfático que se observava neste, a falta de flexão de caso é corrente no Brasil. Câmara Jr., então, busca na "estrutura gramatical do português" as razões para o pronome de 3ª pessoa estar-se "dissociando do sistema casual dos pronomes pessoais", o que seria uma "inovação brasileira". A razão para isso estaria nas diferenças semânticas e formais do pronome de 3ª pessoa em relação aos pronomes de 1ª e 2ª pessoa. Por ter flexão de gênero e número, como os demonstrativos e nomes (ao contrário de *eu* e de *tu*, que "têm plurais heteronímicos") e por estar "ligado aos nomes, dos quais ele é um substituto" (ao contrário dos pronomes de 1ª e 2ª pessoa que "se referem diretamente às pessoas do discurso"), o pronome *ele* se estaria integrando ao quadro dos pronomes demonstrativos e dos nomes, em "um quadro gramatical muito harmônico". E, como os seus novos pares, não exibiria flexão de caso.

Mas esse quadro estrutural não seria suficiente para explicar o processo de variação, pois o pronome de 3ª pessoa tem as mesmas características em Portugal,

e as formas do clítico acusativo estão "bem vivas, no português europeu, desde os primeiros séculos até hoje". Portanto, a inovação brasileira decorre de um "fato fonético que consiste em dois fatos fonéticos entrelaçados" (CÂMARA JR., 1972, p.50): a tendência à próclise sistemática no Brasil, associada à intensificação da primeira consoante do vocábulo fonético. A fragilidade do clítico acusativo da 3ª pessoa decorreria exatamente do fato de ser constituído apenas por uma vogal, ao contrário dos demais pronomes objeto que apresentam o padrão silábico CV(C) (e.g., *me*, *te*, *se*, *lhe* e *nos*) e que por isso se mantêm com grande vitalidade no Brasil, normalmente proclíticos à forma verbal. Portanto, o único pronome pessoal canônico que estaria perdendo a flexão de caso seria o da 3ª pessoa, em razão de sua situação especial no paradigma dos pronomes pessoais. Os demais pronomes canônicos que se mantêm em uso, basicamente os da 1ª pessoa (*eu* e *nós*), não exibiriam variação em sua flexão de caso.

Portanto, pode-se concluir que a variação atual na flexão de caso dos pronomes pessoais na norma culta brasileira tem uma base discursiva e estrutural paradigmática. O fato que provocou a maior desestruturação do sistema antigo foram mudanças de natureza discursiva: a substituição dos pronomes pessoais *tu* e *vós* pelo pronome de tratamento *você(s)* e a substituição do pronome de 1ª pessoa do plural *nós* pela expressão nominal *a gente*, gramaticalizada com essa função. Já fatores estruturais associados a mudanças fonéticas inerentes à formação da prosódia brasileira teriam determinado a perda do clítico acusativo da 3ª pessoa, fazendo com que o pronome *ele* (*ela/eles/elas*), assim como os nomes e demonstrativos aos quais se assemelha, fosse usado invariavelmente em todas as funções sintáticas. Porém, a análise do sistema de pronomes na fala da grande maioria da população brasileira, constituída majoritariamente por indivíduos mestiços e de pouca escolaridade, revelará processos bem mais amplos de variação, que implodem esses esquemas estruturais e funcionais, produzindo formas estranhas ao uso culto e até mesmo rejeitadas pelos falantes com maior nível de escolaridade. Essas formas, ainda presentes na norma popular brasileira, como resquícios de sua origem plurilíngue (LUCCHESI, 2001; MATTOS E SILVA, 2004), só podem ser adequadamente explicadas como o resultado de processos de variação e mudança induzidos pelo contato entre línguas, semelhantes aos que concorreram para a formação das línguas crioulas de base lexical portuguesa da África e da Ásia, não obstante tenham operado em um nível de radicalidade menor do que estes, o que pode ser compreendido a partir do conceito de *transmissão linguística irregular* (cf. capítulo 3 deste livro). Dessa forma, a descrição de como as mudanças desencadeadas pelo contato linguístico massivo, radical e abrupto afetaram a flexão de caso nos crioulos de base lexical portuguesa da África pode ser de grande valia para compreender o processo histórico de formação da norma popular brasileira.

20.2 A flexão de caso dos pronomes pessoais nos crioulos de base lexical portuguesa da África

Uma das características mais generalizadas entre as línguas *pidgins* e crioulas é a perda da flexão de caso dos pronomes pessoais. A eliminação da flexão pronominal de caso presente na língua-alvo que fornece a base lexical para a formação da língua *pidgin* ou crioula integra um amplo processo de simplificação morfológica que compreende também a eliminação da flexão de gênero e de número de determinantes e nomes, e da flexão de número e pessoa dos verbos (cf. capítulo 3 deste livro). Com isso, é normal nas línguas crioulas que a mesma forma do pronome pessoal desempenhe, tanto a função de sujeito, quanto a de complemento verbal ou adjunto adverbial. Nos processos de crioulização do português na África, a flexão de caso dos pronomes pessoais presente na língua-alvo foi, no geral, eliminada, independentemente da radicalidade do processo de crioulização em causa, como se pode ver nos casos do crioulo cabo-verdiano, do crioulo da Guiné-Bissau e do crioulo de São Tomé, apresentados a seguir.

Comparado com as demais línguas crioulas de base lexical portuguesa da África, o crioulo do arquipélago de Cabo Verde (um conjunto de dez ilhas da costa ocidental da África que se posicionam na direção do litoral do Senegal e da Mauritânia) é a língua crioula que mais se aproxima da sua língua-alvo. Mesmo assim, a flexão de caso dos pronomes pessoais do português foi praticamente eliminada nos dois grandes dialetos do crioulo cabo-verdiano (o barlavento e o sotavento), como se pode ver no Quadro 3.

PESSOA DO DISCURSO	FUNÇÃO SINTÁTICA	
	sujeito	complementos e adjuntos verbais
1ª pessoa do singular	*mi* ~ *n*	*me* ~ *mi* ~ *'m*
2ª pessoa do singular	*bô* ~ *bó* ~ *bu*	*bô* ~ *bu* ~ *b'*
3ª pessoa do singular	*êl* ~ *ê*	*l* ~ *êl*
1ª pessoa do plural	*no* ~ *nu* ~ *nos*	*nos* ~ *nus* ~ *n'*
2ª pessoa do plural	*bosês* ~ *osês* ~ *sés*	*bô*
3ª pessoa do plural	*ês*	*ês* ~ *'z*

Quadro 3 - Os pronomes pessoais do crioulo português de Cabo Verde
Fonte: Adaptado a partir de Almada, 1961, p.94-101 e Lopes da Silva, 1984, p.132-133 e 163.

As alterações que se verificam entre as formas apresentadas no Quadro 3 não são de natureza morfológica (o que caracterizaria a flexão casual), mas sim de natureza fônica, decorrentes do processo de cliticização das formas pronominais, como se pode ver nos exemplos abaixo, retirados de Cardoso (1990, p.102):

(8) a. **Mi** e fliz. 'eu sou feliz'

b. **N** tá faze izarsísi. 'eu estou fazendo o exercício'

c. Êl dà-'**m** un livr. 'ele me deu um livro'

(9) a. **Bô** tá faze izarsísi. 'você faz o exercício'.

b. Êl dà-**b**' un livr. 'ele deu a você um livro'

(10) Bo dà-**l** un livr. 'você deu a ele um livro'

(11) a. **No** tá faze izarsísi. 'nós estamos fazendo o exercício'

b. El ojò-**n**'. 'ele nos viu'

O pronome da 1ª pessoa do singular *mi* no exemplo (8a) reduz-se a um travamento nasal quando proclítico ao verbo, em (8b), na função de sujeito, e quando enclítico ao verbo na função de OI, em (8c). O mesmo processo de redução acomete o pronome de 2ª pessoa do singular *bô* na função de OI, em (9b), como também a forma *l*, em (10) (em contraste com *êl*, na função de sujeito, em (9b)) e o pronome *no* em (11b). Assim, pode-se dizer que, no crioulo cabo-verdiano, não há uma flexão morfológica que marque a função sintática dos pronomes pessoais. Deve-se registrar, apenas, uma distinção para as formas tônicas que funcionam, via de regra, como tópicos frasais, como exemplificado em 12:

(12) a. **Ami** e k'el da kel livri. 'a mim é que ele deu o livro'

b. **Abô** e k'el da kel livri. 'a você é que ele deu o livro'

c. **Êl** e k'bo da kel livri. 'a ele é que você deu o livro'

O crioulo português da Guiné-Bissau, apesar de ser um pouco mais distante do português do que o seu congênere de Cabo Verde em vários aspectos gramaticais, exibe uma flexão de gênero mais visível do que esse, como se pode ver no Quadro 4.

PESSOA DO DISCURSO	FUNÇÃO SINTÁTICA		
	sujeito	OD e OI	complemento oblíquo
1ª pessoa singular	*n*	*n*	*mi*
2ª pessoa singular	*bu*	*u*	*bo*
3ª pessoa singular	*i*	*l*	*el*
1ª pessoa plural	*no*	*no*	*nos*
2ª pessoa plural	*bo*	*bo*	*bos*
3ª pessoa plural	*e*	*elis*	*elis*

Quadro 4 - Os pronomes pessoais do crioulo português da Guiné-Bissau
Fonte: Adaptado a partir de Kihm, 1994, p.141-142 e 150-160.

A maior diferenciação se dá nas formas oblíquas. Essas formas são regidas por uma preposição e podem aparecer em posição não argumental como sujeito topicalizado e retomado pela forma átona, como no exemplo abaixo (cf. KIHM, 1994, p.150-151):

(13) **Ami**, **n** ta cria limarya. 'eu, eu crio animais'

Entretanto, entre as funções de sujeito e complementos verbais, praticamente não há diferenciação mórfica que caracterize a flexão de caso, mas alterações fonológicas decorrentes do processo de cliticização dos pronomes, como visto no crioulo cabo-verdiano.

No caso do crioulo da Ilha de São Tomé, que apresenta o maior grau de reestruturação gramatical dentre os três crioulos aqui examinados, o grau de redução morfológica é bem amplo, tendo-se praticamente eliminado a flexão de caso dos pronomes pessoais, como mostra o Quadro 5.

PESSOA DO DISCURSO	FUNÇÃO SINTÁTICA	
	sujeito	objeto
1ª pessoa singular	*n ~ i*^{N}	*mu*^{n} *~ mu ~ m*
2ª pessoa singular	*bo*	*Bo*
3ª pessoa singular	*e ~ ele*	*e ~ ele*
1ª pessoa plural	*nõ*	*Nõ*
2ª pessoa plural	*nãse*	*Nãse*
3ª pessoa plural	*iné*^{N} *~ né*^{N} *~ inê*	*iné*^{n} *~ né*^{n} *~ inê*

Quadro 5 - Os pronomes pessoais do crioulo português de São Tomé

Fonte: Ferraz, 1979, p.62.

A forma do pronome de 1ª pessoa, na posição de sujeito, reduz-se, em função da sua cliticização, a um travamento nasal /N/, cuja realização é determinada pelo ponto de articulação da consoante inicial do verbo: *n-tlaba* 'eu trabalhei'; *m-bi* 'eu vim'; *ng-gosta* 'eu gostei'. Essa redução pode-se dar também em função da vogal nasal *i*^{N}, em todos os contextos: *i*^{n}*-sa* 'eu sou'; *i*^{m}*-tava* 'eu estava'; *i*^{ng}*-kontla* 'eu encontrei' (FERRAZ, 1979, p.62-63).

As formas da 2ª pessoa do singular e da 1ª e 2ª pessoas do plural são invariáveis, sendo sempre expressas pela forma *bo*, *nõ* e *nãse*, respectivamente, em qualquer função sintática. A forma *nãse* tem substituído a forma arcaica *inãse*, podendo ser oxítona — *nãsê* — na posição de complemento. A 3ª pessoa do plural, que é representada por três formas em variação — *iné*^{N} *~ né*^{N} *~ inê* —, também não apresenta flexão de caso (FERRAZ, 1979, p.64-66).

A forma de sujeito da 3ª pessoa do singular sofre uma redução fonológica em decorrência da cliticização na função de sujeito (e.g., *e-fla* 'ele disse'). A forma tônica ocorre em contextos de ênfase: *elê mé sa plumê* 'ele mesmo é o primeiro' (FERRAZ, 1979, p. 64-65).

Em santomense, também se constatam formas tônicas em processos de topicalização do sujeito:

(14) **Ami**, **n**-ga bi. 'eu, eu estou vindo'

Portanto, como no caso de seus dois congêneres, o santomense também eliminou a flexão de caso, subsistindo uma variação fonológica decorrente de processos de cliticização do pronome.

Pode-se concluir, então, que o processo de transmissão linguística irregular mais radical que conduziu à formação dos crioulos portugueses da África implicou a eliminação da flexão de caso dos pronomes pessoais, que apresentam formas átonas, em função da cliticização, e formas tônicas em posição não argumental ou quando regidas por preposição.

Considerando-se que as variedades populares do português brasileiro formaram-se, em grande parte, em contextos sociolinguísticos semelhantes àqueles em que se formaram os crioulos portugueses da África, seria de esperar que processos semelhantes de variação e mudança aos descritos acima também afetassem a flexão de caso dos pronomes pessoais, nas variedades populares do PB. Com efeito, no português popular brasileiro, em geral, e no português afro-brasileiro, em particular, verifica-se um cenário de variação no que concerne à flexão de caso dos pronomes pessoais, em que se podem encontrar formas que apontam para uma simplificação radical da flexão pronominal no processo pretérito de formação dessas variedades linguísticas. Pode-se pensar, dessa forma, que, nos primeiros séculos da formação da sociedade brasileira, até meados do século XIX, houve uma forte erosão do mecanismo da flexão de caso dos pronomes pessoais nas variedades de português falado como segunda língua por escravos africanos e índios aculturados, bem como pelos descendentes desses indivíduos, que já teriam uma variedade do português como língua materna. Essa erosão atingiria todos os pronomes de todas as pessoas do discurso, não se restringindo às formas da 3ª pessoa, pelo caráter nominal desta, nem sendo decorrente de mudanças na pauta dos pronomes pessoais. E o quadro atual de variação no chamado português afro-brasileiro, descrito na próxima seção, confirma essa hipótese, ao demonstrar que se trata de um processo de variação qualitativamente distinto daquele que se observa na norma culta brasileira.

20.3 A flexão de caso dos pronomes pessoais no português afro-brasileiro

Todas as formas de pronomes pessoais nas funções sintáticas de sujeito, OD, OI, complemento oblíquo e adjunto adverbial, depreendidas em um levantamento exaustivo das formas pronominais em 28 entrevistas de tipo sociolinguístico com moradores de quatro comunidades rurais afro-brasileiras isoladas do interior do Estado da Bahia estão reunidas no Quadro 6.[5]

[5] As comunidades estudadas são apresentadas no capítulo 2 deste livro; e a metodologia utilizada na constituição das amostras de fala é descrita no capítulo 5.

PESSOA DO DISCURSO	FUNÇÃO SINTÁTICA			
	sujeito	OD	OI	complemento oblíquo/ adjunto adverbial
1ª pess. sing.	*eu*	***me*** *~ eu*	***me*** *~ para* ***mim***	*de/ni/pra* ***mim*** *~ de/ni eu ~* ***comigo*** *~ mais eu*
2ª pess. sing.	*você ~ tu*	***te*** *~* ***lhe*** *~ você ~ tu*	***te*** *~* ***lhe*** *~ a você*	*pra/com/de você ~ com tu*
3ª pess. sing.	*ele/a*	*ele/a*	*(pra) ele/a*	*de/pra/ni/com ~ mais ele/a*
1ª pessoa pl.	*nós ~ a gente*	*a gente ~ nós*	*pra/a a gente ~ nós*	*pra/de/com ~ mais a gente ~ nós*
2ª pessoa pl.	*vocês*	*vocês*	*pra vocês*	*pra/mais vocês*
3ª pessoa pl.	*eles/as*	*elas/eles*	*pra(a) eles*	*de/ni/pra/com ~mais eles/as*

Quadro 6 - Os pronomes pessoais no português afro-brasileiro

As formas pronominais flexionadas em caso para as funções sintáticas de complementos verbais e adjuntos adverbiais (grifadas em negrito no Quadro 6) restringem-se, na gramática do português afro-brasileiro, à 1ª e à 2ª pessoas do singular. Além disso, o uso das formas do chamado caso reto *eu*, *tu* e *nós* em contextos em que na norma culta ocorrem as formas do caso oblíquo demonstra que a perda da flexão de caso no português afro-brasileiro se deu independentemente da introdução das formas *você(s)* e *a gente*. O comentário sobre cada pessoa do discurso pode explicitar isso.

Na 1ª pessoa do singular, a forma do caso reto encontra-se em variação com as formas oblíquas nas funções de: OD (e.g., *meu pai* ***me*** *criô / o marido num quis* ***eu*** *não*), complemento oblíquo (e.g., *todo mundo depende de* ***mim*** */ eles depende de* ***eu***) e adjunto adverbial (e.g., *tu vem ficá aqui* ***comigo*** */ ela num quis ir mais* ***eu*** *não*). Em seu conjunto, a frequência de formas não padrão fica em torno de 6% do total de ocorrências do pronome de 1ª pessoa do singular nas funções de OD, complemento oblíquo e adjunto adverbial (20 ocorrências não padrão em um total de 318 ocorrências). A forma do caso reto *eu* não se documenta na amostra de fala analisada na função de OI, registrando-se apenas as formas ***me*** e ***mim*** (e.g., ***me*** *dá meu 'menduim; deu pra* ***mim***). Entretanto, apesar de não ter se constatado na amostra de fala analisada, pode-se afirmar que uma frase como *ela entregou pra eu* é possível, tanto na fala afro-brasileira, quanto na fala popular rural. Portanto, as formas flexionadas em caso da 1ª pessoa do singular (*me, mim* e *comigo*) sofrem, na fala das comunidades rurais afro-brasileiras, a variação com a forma não marcada do caso reto, em todas as posições sintáticas, mesmo que o nível dessa variação seja baixo nos dias atuais.

A introdução de *você*, e suas formas reduzidas *ocê* e *cê*, como pronome de 2ª pessoa implica o seu uso, como forma invariável quanto ao caso, nas posições de complemento verbal e adjunto adverbial: *eu vi* ***ocê****; eu cansei de falá com* ***você****; vê se o menino tira uns côcos p'****ocê***. Ainda como decorrência da introdução de *você*, usa-se *lhe*, com referência à 2ª pessoa, tanto na função canônica de OI, quanto no uso não

padrão como OD: *eu vô **lhe** contá* (OI); *eu já **lhe** vi* (OD). Esses fatos, observados na amostra de fala analisada, são comuns a todas as variedades do PB, incluindo a sua norma urbana culta. Entretanto, alguns fatos do português afro-brasileiro com relação à flexão de caso com a 2ª pessoa do singular, que podem ocorrer em outras variedades populares e rurais do PB, não se verificam na norma culta. A manutenção de *tu* em variação com a forma inovadora *você* na função de sujeito está relacionada com o uso da forma oblíqua *te*, para as funções de OD e OI: *eu vô **te** levá po Cinzento*; *não **te** dô a conta*, mas não se registraram as formas tônicas *ti* e *contigo*. Por outro lado, o uso do pronome do caso reto *tu* na função de OD e complemento oblíquo (e.g., *só num já matei **tu***, *num falei com **tu***), encontrado na amostra, aponta para um quadro de variação no português afro-brasileiro, e em muitas variedades populares do PB, que se relaciona com a mudança crioulizante de eliminação da flexão de caso dos pronomes pessoais. Por fim, vale registrar que, como verificado em relação à 1ª pessoa, não foram observadas, na amostra analisada, realizações de *tu* na função de OI, embora se possa dizer que tal uso é bastante factível na fala popular (e.g., *ela não deu o livro **pra tu** não*).

Na 3ª pessoa do singular e do plural, não se registraram, na amostra de fala analisada, formas oblíquas, observando-se exclusivamente o uso dos pronomes do caso reto em todas as funções sintáticas. Porém, diferentemente do que se disse com relação ao uso de *eu* e de *tu* fora da posição de sujeito, o uso de *ele* como acusativo é geral no PB (cf. seção 20.1 deste capítulo). Na amostra analisada também não ocorreram *lhe*(*s*) com referência à 3ª pessoa, nem as formas tônicas *si* e *consigo*, podendo-se dizer que, no português afro-brasileiro, a flexão de caso foi eliminada na referência à 3ª pessoa.

A gramaticalização de *a gente* como pronome pessoal de 1ª pessoa do plural no português brasileiro figura na função de sujeito e em todas as demais funções sintáticas. O fato divergente que se constata na amostra analisada é o uso da forma do caso reto do pronome canônico *nós* nas posições de OD e adjunto adverbial (e.g., *o padre Joselito levô **nós**...*; e *ele é nascido e criado **mais nós** aí no Cinzento*, respectivamente), enquanto na norma culta essas funções são desempenhadas pelas formas oblíquas *nos* e *conosco*. Além disso, documentou-se também a forma *nós* não flexionada na função de OI: *porque é uma coisa que Deus deixô, Deus deu **nós**, né*, e *vendêro **nós** esta terra bruta e deu um recibo*; o que não havia sido observado em relação à 1ª e à 2ª pessoas do singular. Esse uso divergente de *nós* demonstra, também na 1ª pessoa do plural, que a perda da flexão de caso no português popular não se deu em função de substituições na pauta dos pronomes pessoais, pois o uso da mesma forma para as diversas funções sintáticas não se dá apenas com a forma nominal *a gente*, mas também com a forma pronominal *nós*, do mesmo modo que com os pronomes *tu* e *eu*. Resta inexoravelmente a explicação de que esse quadro de variação decorre de mudanças pretéritas induzidas pelo contato entre

línguas, já que o paralelo com os processos de mudança constatados nos crioulos africanos de base lexical portuguesa é evidente.

Na 2ª pessoa do plural, representada apenas pelo pronome *vocês*, a flexão de caso também está ausente, não se encontrando qualquer ocorrência de formas oblíquas, como *lhes*, *si* e *consigo*, presentes na norma culta, em frases como: *eu já* ***lhes*** *disse que vocês devem se comportar*; *vocês só dependem de* ***si***; *vocês só podem contar* ***consigo***.

Portanto, a flexão de caso dos pronomes pessoais no português afro-brasileiro reduz-se atualmente à 1ª e à 2ª pessoas do singular. E, mesmo nesses casos, observa-se um quadro de variação entre formas flexionadas e formas não marcadas do caso reto. O uso de *eu*, de *tu* e de *nós* nas funções de complemento verbal e adjunto adverbial aponta para um estágio anterior no processo de formação dessas comunidades de fala, em que a variação na flexão de caso dos pronomes pessoais atingia maciçamente todas as pessoas do discurso, mesmo quando essas pessoas eram representadas pelas formas pronominais canônicas (*eu*, *tu* e *nós*), antes da introdução de *você* e de *a gente*. Esse tipo de variação é característico dos processos de transmissão linguística irregular que se desencadeiam em situações de contato entre línguas e marcaram a constituição das variedades populares do PB, sobretudo no caso das comunidades rurais afro-brasileiras, que se formaram a partir de quilombos, ou em torno de grandes propriedades rurais do interior do país, e que concentraram um largo contingente de mão de obra escrava. A recomposição da flexão de caso, sobretudo em relação à 1ª pessoa do singular, deve ser vista como um processo mais recente de mudança do tipo descrioulizante, que resulta de influências externas crescentes a partir da intensificação do contato entre essas comunidades e os demais segmentos da sociedade brasileira, sobretudo a partir da segunda metade do século XX (cf. capítulo 1 deste livro).

Deve-se, então, estabelecer uma diferença meramente quantitativa entre os processos de mudança que atingiram os crioulos portugueses da África e o português afro-brasileiro. Enquanto naqueles, um processo de transmissão linguística radical levou à eliminação completa da flexão de caso dos pronomes pessoais; no caso do português afro-brasileiro, um processo de transmissão linguística irregular mais leve produziu apenas um quadro de variação em que as formas não marcadas do caso reto coocorrem com as formas flexionadas do caso oblíquo nas funções de complemento verbal e adjunto adverbial. Além disso, também não se observa, no português afro-brasileiro, o intenso processo de cliticização dos pronomes pessoais, principalmente na posição de sujeito, que se documenta em todos os crioulos portugueses da África, e que deu azo ao surgimento de formas tônicas diferenciadas, usadas em posições não argumentais.

Por outro lado, constata-se a conjugação de processos de natureza crioulizante no português brasileiro e no português popular como um todo, como o uso

do advérbio *mais* com o valor da preposição *com*; uso esse que se combina apenas com as formas não flexionadas do caso reto (*Joselito faz a missa aqui* ***mais eu***), sendo agramatical a combinação com as formas flexionadas do caso oblíquo (a frase *Joselito faz a missa aqui* ***mais mim*** seria agramatical). A substituição da preposição *com* pelo advérbio *mais* conforma-se ao princípio da **transparência semântica**, invocado para explicar muitas mudanças que caracterizam a formação das línguas *pidgins* e crioulas (SEUREN; WEKKER, 1986). Segundo esse princípio, nas situações de contato maciço entre línguas, as formas de significado mais transparente tendem a sobrepujar aquelas formas mais gramaticais de decodificação mais difícil. Além disso, tal processo, que deve ser definido como um processo de gramaticalização do advérbio *mais* como preposição, ajusta-se perfeitamente ao cenário de formação das línguas *pidgins* e crioulas, em que formas lexicais são requeridas para ocupar as lacunas deixadas pelo desaparecimento das partículas gramaticais da língua-alvo (BICKERTON, 1999; e capítulo 3 deste livro). Por outro lado, a agramaticalidade da combinação entre o pronome oblíquo e a forma *mais* na função de preposição ajusta-se ao princípio da coesão estrutural proposto por Lucchesi (2000), que prediz a coocorrência de estruturas crioulizantes na mesma porção da sentença (cf. seção 4.2.2 do quarto capítulo deste livro).

Outra estrutura de característica crioulizante é a combinação da forma *ni* da preposição, usada como preposição locativa multifuncional, com as formas do caso reto: *ele vem botá isso* ***ni eu***. Porém, nesse caso, o princípio da coesão estrutural não é aplicado de forma categórica, podendo a forma *ni* se combinar com a forma oblíqua: *ele vem botá isso* ***ni mim***.

20.4 Conclusão

Os fatos aqui apresentados com referência à flexão de caso dos pronomes pessoais comprovam mais uma vez a polarização da realidade sociolinguística brasileira, nos termos da sistematização de Lucchesi (2001, 2002a, 2002b, 2006a).[6] Além disso, revelam que a polarização atual resulta de processos históricos de formação da norma culta e da norma popular, que são distintos e independentes, não obstante a mútua influência que, por suposto, existiu, desde o início da colonização do Brasil, entre os segmentos da elite e da base da sociedade brasileira, em maior ou menor grau, a depender do contexto e da época que se considere.

Esses processos de evolução distintos manifestam-se claramente em relação ao mecanismo da flexão de caso dos pronomes pessoais. Por um lado, na formação

[6] Cf. capítulo 1 deste livro.

da norma culta brasileira, ocorreram mudanças que estão diretamente ligadas a inovações discursivas e pressões estruturais. Essas mudanças, que vão definir os contornos atuais da norma culta brasileira, se teriam acelerado a partir do século XX, com a diminuição da influência do modelo europeu, não obstante o esforço da tradição gramatical brasileira em conservar o padrão lusitano como referência para a normatização linguística (LUCCHESI, 2001; 2002a). Por outro lado, constata-se, na história das variedades populares do PB, a erosão da morfologia flexional em função da precária aquisição do português como segunda língua por parte dos índios que iam sendo aculturados e dos africanos que eram trazidos para o Brasil como escravos. Durante a socialização e a nativização desse modelo defectivo de segunda língua entre os índios e os afrodescendentes, a expressão morfológica dos mecanismos gramaticais afetados se foi recompondo, mas os reflexos da erosão inicial se fazem presentes até os dias de hoje, embora, em alguns aspectos, de forma bastante reduzida, quase residual.

Desse modo, o contraste entre as projeções diacrônicas feitas em vários capítulos que compõem a segunda parte deste livro a partir do estado atual do português afro-brasileiro e os dados disponíveis sobre a formação da norma culta das grandes cidades do país tem revelado uma clivagem histórica, cujo resultado é o quadro atual de polarização sociolinguística do Brasil, na medida em que a diferença entre os quadros de variação nesses dois extremos da realidade sociolinguística brasileira só podem ser compreendidos satisfatoriamente, no plano de sua constituição histórica, quando se postulam processos distintos e independentes de mudança linguística, embora não se possa desconsiderar a sua mútua influência. Tal diferença radica no fato de que, enquanto a relação com os processos sociais é mais direta na formação do português afro-brasileiro, os processos de mudança na norma urbana culta são mais restringidos por fatores estruturais. Não se trata, porém, nesse último caso, de pensar que a mudança linguística resulte exclusivamente do jogo cambiante de elementos da estrutura linguística operando no vácuo social, até porque a mudança linguística não pode ser pensada seriamente somente nesses termos.

No caso específico da perda da flexão de caso dos pronomes pessoais na norma culta, identifica-se, por um lado, como móvel da mudança a substituição de itens da pauta dos pronomes sociais: *tu* e *vós* por *você(s)*, e *nós* pela forma *a gente*. Tal mudança, que tem uma natureza quase lexical de substituição na designação de uma pessoa do discurso (razão pela qual foi aqui denominada de *mudança discursiva*), certamente está correlacionada a mudanças nas relações sociais, cujo reflexo seria a alteração na forma como os falantes se referem aos participantes da interação verbal. Mas não deixa de ser uma mudança no interior de uma comunidade de fala, sem a intervenção de um grupo adventício, falante de outra(s) língua(s). E é essa mudança discursiva de raiz social que vai constituir um dos

dois grandes fatores de desequilíbrio do sistema de flexão de caso na norma urbana culta brasileira. O outro fator seria a alteração na prosódia do português no Brasil com fortalecimento da pronúncia das vogais átonas; processo fônico, cujas origens sociais ainda estão por ser desvendadas, não se podendo descartar a possibilidade de estar associado às situações de multilinguismo do Brasil nos séculos XVII e XVIII. Tal mudança prosódica teria comprometido irremediavelmente o uso do clítico acusativo de 3ª pessoa, desencadeando um processo de mudança em cadeia que estaria conduzindo ao uso do pronome *ele* como um pronome invariável quanto ao caso; mudança essa que seria favorecida, por um lado, pelo caráter nominal desse pronome (com flexão de gênero e número, como os nomes) e pelo emprego da forma de dativo *lhe*(*s*) como referência preferencial à 2ª pessoa do discurso. Portanto, as origens discursivas, bem como as restrições estruturais, acabaram por circunscrever, na norma culta, a falta de flexão de caso às formas de origem nominal, *você*(*s*) e *a gente*, e ao pronome de 3ª pessoa, mantendo-se intacta a flexão de caso para as formas essencialmente nominais, nomeadamente às da 1ª pessoa.

O fato de a variação na flexão casual afetar as formas genuinamente pronominais no português afro-brasileiro (e em grande parte do português popular do Brasil) coloca de forma irrefutável a necessidade de se pensar em um processo de mudança linguística com determinações sociais muito mais amplas e profundas. Desse modo, o paralelo com os processos de contato entre línguas massivo, radical e abrupto, em geral, e com crioulização do português na África, em particular, impõe-se inexoravelmente. A erosão da flexão de caso é uma mudança característica das situações de contato entre línguas de que resulta uma variedade linguística historicamente original (um *pidgin* ou crioulo). Os crioulos de base portuguesa da África não fogem a essa regra, exibindo um quadro muito reduzido de flexão pronominal de caso, muitas vezes restrita a alterações fônicas decorrentes da cliticização dos pronomes pessoais.

Como na África, os escravos africanos trazidos para o Brasil também aprenderam precariamente o português, e esse modelo defectivo de segunda língua serviu como modelo para aquisição da língua materna dos seus descendentes. É natural, portanto, que o mesmo processo que lá afetou a flexão de caso dos pronomes pessoais tenha ocorrido também aqui, porém com uma intensidade menor, pois as condições sócio-históricas diferenciadas do Brasil permitiram um maior acesso dos afrodescendentes aos modelos da língua-alvo, impedindo que esse mecanismo gramatical fosse totalmente eliminado, mas produzindo um quadro de variação mais ou menos intenso em cada caso, consoante o grau de radicalidade do contato linguístico (cf. capítulos 1, 3 e 4, deste livro). Em variedades crioulizadas do português que eventualmente tenham se formado em torno de grandes engenhos e quilombos no interior recôndito do Brasil, durante os séculos XVII, XVIII e XIX, a flexão de caso dos pronomes pessoais pode ter sido muito reduzida. Já nas pequenas

propriedades rurais, com um pequeno número de escravos, e na periferia das vilas e cidades deve ter-se originado um quadro de variação um pouco mais intenso do que o que se observa hoje nas comunidades afro-brasileiras isoladas e em outras variedades rurais. Essa variação que afetaria todos os pronomes pessoais foi sendo progressivamente eliminada, na medida em que aumentava o acesso dos afrodescendentes e mestiços aos modelos da língua-alvo, introduzindo-se inicialmente as formas oblíquas que se tinham mantido incólumes na fala culta, como as da 1ª pessoa do singular, não atingindo, entretanto, as formas que já haviam caído em desuso, em função da introdução de *você* e de *a gente* e daquelas da 3ª pessoa em que os fatores estruturais favoráveis teriam permitido a sua eliminação.[7] O resultado desse processo seria o quadro atual, em que uma pequena variação afetando a flexão de caso das formas pronominais por excelência (*eu*, *tu* e *nós*) se combina com a eliminação da flexão de caso no pronome de 3ª pessoa, que constitui o elo fraco da cadeia. Assim sendo, apesar da grande diferença na intensidade da erosão desse mecanismo gramatical que separa o português afro-brasileiro dos crioulos portugueses da África, essa diferença seria meramente quantitativa e não qualitativa, porquanto o processo seminal de mudança nos dois casos seria o mesmo.

Se, para o português afro-brasileiro e para grande parte do português popular do Brasil, o quadro atual de variação remonta a processos de mudança decorrentes da interação maciça de povos falantes de línguas diversas, a variação na norma culta brasileira, bem como a recomposição da flexão de caso nas variedades populares estaria relacionada historicamente a uma dialética da interação de segmentos distintos na formação da sociedade brasileira. Nesse sentido, o contato dialetal pode ser invocado para explicar um processo de mútua influência. Se os falantes das variedades populares do PB tinham e têm como modelo o padrão de fala dos segmentos mais altos da sociedade brasileira, o padrão dito culto também tem sido afetado pelo convívio com os segmentos populares desde os primeiros séculos da colonização do Brasil. Os filhos dos senhores de engenho conviviam com amas de leite e filhos de escravos, que tinham na sua fala as marcas do processo de transmissão linguística irregular acima descrito. Desse modo, o modelo de português dos seus pais, tios e avós (muitas vezes portugueses) sofria a concorrência dos modelos alterados dos escravos com que conviviam. O poder e o prestígio faziam a balança pender para o modelo lusitanizante, mas a convivência com a fala "estropiada" dos escravos não deixaria de "contaminar" o seu comportamento linguístico, sobretudo nos casos em que as condições estruturais fossem mais favoráveis. Desse modo, pode-se identificar o móvel social das mudanças discursivas

[7] Todo esse raciocínio se aplica também aos índios aculturados e mamelucos, nas regiões de grande concentração indígena, e onde esses índios não tenham sido dizimados.

que modificaram o elenco dos pronomes pessoais em todo o português brasileiro, bem como da alteração na prosódia brasileira que está na origem da perda de flexão de caso do pronome de 3ª pessoa.

Portanto, esta análise da variação atual na flexão de caso fundada no cotejo de dois polos opostos do universo sociolinguístico brasileiro acabou por revelar como processos históricos distintos se conjugaram dialeticamente para dar forma à atual clivagem sociolinguística do Brasil.

21 O sistema de expressão de posse

Dante Lucchesi
Silvana Silva de Farias Araújo

A expressão da posse no português do Brasil por meio dos pronomes possessivos tem sido tema de muitas investigações, a exemplo dos trabalhos de Silva (1982, 1984, 1996), Perini (1985), Kato (1985), Neves (1993, 2000), Cerqueira (1993, 1996) e Muller (1997). Esse interesse é provocado por uma série de mudanças que se insinuam mesmo nos padrões de comportamento linguístico dos segmentos sociais brasileiros de maior nível de escolarização. Entre essas, pode-se destacar a ascensão do pronome *você* como pronome de referência ao interlocutor na maior parte do território brasileiro, o que implicou no uso extensivo do possessivo *seu* (e suas formas flexionadas: *sua*, *seus*, *suas*) com referência à segunda pessoa do discurso. Disso resulta o crescente uso do Sintagma Preposicionado (SP) *dele* (e suas formas flexionadas: *dela*, *deles*, *delas*) para expressar sem ambiguidade a posse com referência à terceira pessoa do discurso. Do mesmo modo, a gramaticalização da expressão nominal *a gente* como forma variante do pronome de 1ª pessoa do plural (LOPES, 1999) atinge também a sua expressão de posse, com variação entre o pronome canônico *nosso* (e suas formas flexionadas: *nossa*, *nossos* e *nossas*) e a expressão *da gente*, expressão analítica e inflexionável quanto ao gênero e ao número.

Este capítulo focaliza a expressão de posse na gramática do português afro-brasileiro. Para além dos processos de variação referidos acima, observa-se também na fala das comunidades rurais afro-brasileiras isoladas outros processos de variação que podem ser relacionados mais diretamente com as situações de contato entre línguas. Tais processos, que estão presentes nas variedades rurais e populares do português brasileiro (PB), são mais notáveis nessas comunidades em função das especificidades de sua formação histórica.[1] De acordo com a fundamentação teórica apresentada no capítulo 3, as variedades linguísticas formadas no processo de *transmissão linguística irregular* se caracterizam por uma perda de morfologia flexional em relação à língua-alvo e pela substituição de formas sintéticas por expressões analíticas. A flexão de caso, que compreende o caso morfológico dos pronomes

[1] Ver o capítulo 2 deste livro.

possessivos, é um dos itens que via de regra não se conservam nas línguas *pidgins* e crioulas, em que uma única forma pronominal costuma desempenhar tanto a função de sujeito, quanto a de objeto e a de genitivo (cf. capítulo anterior). Desse modo, expressões como *é de eu* (em vez de *é meu*) e *terra de nós* (em vez de *nossa terra*), que ainda podem ser encontradas nas variedades rurais mais recônditas do português brasileiro, podem ser vistas como vestígios de processos de alterações mais amplos que teriam caracterizado as antigas variedades linguísticas que se formaram na socialização e nativização de um modelo defectivo de língua portuguesa adquirida precariamente como segunda língua por uma ampla massa de escravos africanos e indígenas aculturados, no interior do país, no período da Colônia e do Império.

Com base nessa visão, foi feita uma descrição do sistema de posse no português afro-brasileiro, procurando identificar as formas divergentes em relação ao uso culto que refletiriam processos pretéritos de mudanças induzidas pelo contato entre línguas. Além disso, fez-se uma análise mais vertical, focalizando a variação na expressão da posse com referência à 1ª pessoa do plural, com base nos pressupostos teórico-metodológicos da Sociolinguística Quantitativa (WEINREICH; LABOV; HERZOG, 1968 [2006]; LABOV, 1972a, 1982, 1994). Na expressão de posse referente ao *eu-ampliado* (BENVENISTE, 1982), foram registradas no *corpus* analisado três formas variantes: a forma padrão *nosso* (e suas formas flexionadas: *nossa*, *nossos* e *nossas*) e duas formas analíticas: *da gente* e *de nós*. A primeira dessas formas analíticas é documentada em variedades urbanas do português do Brasil, mas com baixa frequência de uso, conforme atestam os estudos de Silva (1982, 1984, 1996), Neves (1993, 2000) e Cerqueira (1993, 1996), dentre outros, enquanto a segunda é considerada neste estudo como um vestígio do contato entre línguas. Assim, este capítulo está estruturado da seguinte maneira: (i) uma informação sobre os pronomes possessivos nas línguas crioulas; (ii) uma revisão da literatura sobre o sistema de posse na língua portuguesa; (iii) uma descrição do sistema de posse encontrado nas comunidades analisadas; (iv) uma análise sociolinguística da variação na expressão de posse com referência à 1ª pessoa do plural; e as suas conclusões.

21.1 Os pronomes possessivos nas línguas crioulas

A perda da flexão de caso dos pronomes é uma das características gerais das línguas *pidgins* e crioulas (cf. capítulo anterior). Tal mudança ocorre no bojo do amplo processo de redução da morfologia aparente, inerente aos processos de **transmissão linguística irregular** (cf. capítulo 3 deste livro). Com isso, é normal

nas línguas crioulas que a mesma forma do pronome pessoal desempenhe tanto a função de sujeito quanto a de objeto, bem como a função de expressar a posse (caso genitivo); eliminando-se habitualmente a flexão de caso presente na língua lexificadora. Os crioulos de base lexical portuguesa da África, em maior ou menor grau, se conformam a essa tendência, em função da intensidade do processo de reestruturação gramatical de cada língua crioula, em seu processo histórico de formação.

O crioulo cabo-verdiano é a língua nacional do Arquipélago de Cabo Verde e apresenta dois grandes dialetos (o barlavento e o sotavento). É a língua crioula de base lexical portuguesa da África que mais se aproxima da língua-alvo, em função de uma maior presença do colonizador português em seu processo histórico de formação, sobretudo se comparada com as línguas crioulas de base lexical portuguesa das Ilhas de São Tomé e Príncipe. Em função dessas características sócio-históricas, conservam-se, na pauta dos pronomes possessivos do cabo-verdiano, algumas formas flexionadas da língua portuguesa, como se pode ver no Quadro 1.

PESSOA	FUNÇÃO SINTÁTICA		
	sujeito	objeto	possessivo
1ª pessoa sing.	mi ~ n	me ~ mi ~ 'm	ña
2ª pessoa sing.	bô ~ bó ~ bu	bô ~ bu ~ b'	bô
3ª pessoa sing.	êl ~ ê	l ~ êl	sê ~ si
1ª pessoa pl.	no ~ nu ~ nos	nos ~ nus	nos
2ª pessoa pl.	bosês ~ osês ~ sês	bô	bosês ~ osês ~ sês
3ª pessoa pl.	ês	ês ~ 'z	sés

Quadro 1 - Os pronomes pessoais do crioulo português de Cabo Verde

Fonte: Adaptação a partir Almada (1961, p.94-101) e Lopes da Silva (1984, p.132-133 e 163).

As formas próprias do possessivo, que guardam, portanto, uma flexão de caso genitivo, são as formas da 1ª pessoa do singular (*ña* < ptg. *minha*) e as da 3ª pessoa do singular e plural (*sê/sês* < ptg. *seu/seus*). As formas da 2ª pessoa e da 1ª pessoa do plural via de regra não se flexionam para expressar a posse. Independentemente dessa possibilidade de flexão, todas essas formas são empregadas adjetivamente e precedem o nome determinado, como em: *ña kaza* 'minha casa' e *bô kaza* 'tua casa'. Lopes da Silva (1984, p.163) menciona uma possível flexão de número, referindo-se às formas *ñas* 'meus, minhas', *bôs* 'teu, tuas' e *sês* 'seus, suas'. Tal flexão poderia ser interpretada como um reflexo da descrioulização. Por outro lado, esse autor afirma também não haver resquícios de flexão de gênero, funcionando todas as formas tanto para o masculino, quanto para o feminino.

Para o emprego como substantivo ou predicativamente, Lopes da Silva (1984) e Almada (1961) referem-se às formas flexionadas *meu*, *miña*, *bósa*, *nósa*, *seu* e *seus*, que podem vir regidas pela preposição ou não, como em:

(1) *ês kaza ê (de) bósa*

'esta casa é tua'

(2) *ês kaza ê (de) miña~meu*

'esta casa é minha'

Essas estruturas com um SP para indicar a posse são sempre usadas quando se emprega um demonstrativo no SN, como em *ês kaza de miña~meu* 'esta minha casa'. Já com o emprego do numeral, mantém-se uma estrutura análoga à da língua-alvo como em *ñas sinku kaza* 'minhas cinco casas' (ALMADA, 1961, p.163).

No crioulo português da Guiné-Bissau, conhecido localmente apenas pela designação *Crioulo*, encontramos uma situação híbrida, que pode ser atribuída ao grau intermediário de reestruturação do Crioulo em face aos seus congêneres africanos. O Quadro 2 reúne os pronomes em questão.

PESSOA	FUNÇÃO SINTÁTICA		
	sujeito	objeto	possessivo
1ª pessoa singular	n	n	nya
2ª pessoa singular	bu	u	su
3ª pessoa singular	i	l	si
1ª pessoa plural	no	no	no
2ª pessoa plural	bo	bo	bo
3ª pessoa plural	e	elis	se[2]

Quadro 2 - Os pronomes pessoais do crioulo português da Guiné-Bissau

Fonte: Adaptação a partir de Kihm (1994, p.141-142 e 150-160).

Como se pode ver, enquanto na 2ª e 3ª pessoas do singular, bem como na 3ª pessoa do plural, existe uma forma diferenciada para o possessivo, o sincretismo é evidente na 1ª e 2ª pessoas do plural. No caso da 1ª pessoa do singular, a diferenciação pode decorrer da aplicação de uma regra morfofonética, não se configurando, assim, uma flexão de caso propriamente dita. No que concerne à ordem, no Crioulo, assim, como no português, o possessivo adjetivo normalmente precede o nome como no exemplo abaixo:

[2] Kihm (1994, p. 142) chama a atenção para o fato de, em Crioulo, se ter desenvolvido uma diferenciação entre a forma do possessivo da 3ª pessoa do singular e do plural, respectivamente *si* e *se*, distinção essa inexistente no português. Mas, segundo o autor, o processo diacrônico dessa bifurcação a partir da forma lexicogênica *seu* do português permanece desconhecido.

(3) *I furtadu si galinya*

‘roubaram sua galinha (dele/dela)’

Já o crioulo português de São Tomé, o *santomense*, ou simplesmente *forro*, exibe estruturas bastante divergentes em relação ao português, por se tratar de uma das línguas de maior grau de reestruturação gramatical no cenário das línguas crioulas de base lexical portuguesa da África. Nesse sentido, o grau de redução morfológica é bem amplo, tendo-se praticamente eliminado a flexão de caso dos pronomes pessoais, como se pode observar no Quadro 3.

PESSOA	FUNÇÃO SINTÁTICA		
	sujeito	objeto	possessivo
1ª pessoa sngular	n ~ i^{N} [3]	mu^{N} ~ mu ~ m	mu^{N} ~ mu ~ m
2ª pessoa sngular	bo	bo	bo
3ª pessoa sngular	e ~ elê	e ~ elê	(d)e
1ª pessoa plural	nõ	nõ	nõ
2ª pessoa plural	nãse	nãse	nãse ~ nãsê
3ª pessoa plural	iné^{N} ~ né^{N} ~ inê	iné^{N} ~ né^{N} ~ inê	iné^{N} ~ né^{N} ~ inê ~ ne

Quadro 3 - Os pronomes pessoais do crioulo português de São Tomé
Fonte: Ferraz, 1979, p.62.

Afora a intensa variação morfofonética das formas, praticamente não se observa flexão de caso nos pronomes pessoais em santomense, o que é igualmente válido para o genitivo. Ferraz (1979, p.61-62) afirma que não há um equivalente para os pronomes possessivos do português (como *meu*, *seu*, *nosso*), já que “em santomense os pronomes objetos são usados para indicar o caso genitivo, como em: *e be mu*^{N} ‘ele viu me’, *lumõ mu*^{N} ‘meu irmão’”. Outro aspecto notavelmente divergente em relação à língua lexificadora é a ordem. Enquanto no português o possessivo geralmente precede o nome, em santomense o possessivo geralmente segue o nome determinado, como em (4) abaixo; podendo, entretanto, preceder o nome, como em (5).

(4) *e sa livlu mu*^{N}

‘é o meu livro’

(5) *ke mu ka fika zuntu d’awa*

‘a minha casa fica junto do rio’

---

[3] O N é um morfema nasal foneticamente dependente que assimila o traço ponto de articulação da consoante do verbo que se segue ao pronome sujeito, como em: *n’tlaba* ‘eu trabalhei’; *m’bi* ‘eu vim’. Essa redução fonológica decorrente da cliticização pode ocorrer também em função da vogal nasal i^{N}, em todos os contextos (FERRAZ, 1979, p. 62-63).

Outro aspecto interessante a ser registrado é o dos SPs regidos pela preposição *di* (< ptg. *de*), que podem assumir um valor genitivo. Nessa estrutura, verifica-se uma regra de apagamento da preposição regida por fatores de fonética sintática, tanto com nomes quanto com pronomes (FERRAZ, 1979, p.69-70). Com nomes iniciados por consoante, a preposição é opcionalmente apagada, como em:

(6) *m'basu (di) pota*

'em baixo da porta'

(7) *zozé sa ta sondu ni tlachi (di) zõ*

'José está sentado atrás de João'

No caso dos pronomes pessoais iniciados com consoante, a preposição *di* é obrigatoriamente apagada, sendo agramatical uma estrutura como **livlu di mu*^{N}, só sendo aceita a expressão *livlu mu*^{N} 'meu livro'. Já com nomes e pronomes iniciados com vogal, a realização da preposição *di* é obrigatória, verificando-se a elisão, como em:

(8) *e sa i basu daluvu*

'ele está de baixo da árvore'

(9) *mina kondê di lumõ de*

'a menina se escondeu do irmão dela'

No caso do apagamento do núcleo nominal, o nome é substituído por um pronome *ji* que se funde com a preposição *di*, quando ela se realiza, como se pode ver no cotejo dos exemplos (10) e (11), com (12) e (13), respectivamente:

(10) *e sa livlu mu*^{N}

'é o meu livro'

(11) *e sa livlu diné*^{N}

'é o livro deles/as'

(12) *e sa ji mu*^{N}

'é o meu'

(13) *e sa jiné*^{N}

'é o deles/as'

Essa breve descrição do sistema de indicação de posse das línguas crioulas de base lexical portuguesa da África pode fornecer subsídios interessantes sobre como o processo de transmissão linguística irregular desencadeado em situações de contato linguístico maciço pode afetar esse aspecto da gramática. O resultado mais radical seria a preservação de uma única forma invariável do pronome pessoal para todas as funções sintáticas, inclusive a sua adjunção ao nome para indicar a

posse, ou seja, o caso genitivo. Paralelamente, registra-se o uso de formas analíticas com a preposição *de*, sobretudo nos casos de apagamento do núcleo nominal, ou quando o pronome é empregado como predicativo do sujeito.

Nas variedades populares do português brasileiro, mesmo nas comunidades rurais afro-brasileiras isoladas, não se observa o resultado mais radical de utilização de uma única forma invariável também com a função de genitivo. Entretanto, verifica-se o emprego de formas analíticas, tais como *de eu* e *de nós*, em vez de *meu/minha/meus/minhas* e *nosso/nossa/nossos/nossas*, respectivamente. Essas formas são bastante divergentes do uso culto brasileiro, realçando uma relação histórica entre elas e os processos de mudança induzidos pelo contato entre línguas, como se buscará demonstrar nas seções seguintes.

21.2 A expressão da posse em português

Em português, a posse pode ser indicada por expressões referenciais, circunscritas à terceira pessoa do discurso, como em *a casa da Maria*, que são introduzidas pela preposição *de*. Tais construções surgem na passagem do latim ao português em função da perda da flexão de caso dos nomes, extinguindo-se, por conseguinte, as expressões sintéticas de posse do tipo *domus Mariae*. Já no caso dos pronomes pessoais, a flexão de caso se conservou no português, de modo que se manteve a forma *minha casa* (< lat. *domum meam*), não permitindo que a forma analítica **casa de eu* se consagrasse na variedade românica que se desenvolveu no Noroeste da península ibérica. Para além da flexão de caso, os pronomes possessivos também se flexionam quanto ao gênero e ao número para concordar com o elemento possuído, como exemplificado em: *minha casa, meu chefe, minhas pernas, meus pensamentos*.

Assim, para a indicação da posse, o português conta com os pronomes possessivos e SPs introduzidos pela preposição *de* que podem ser definidos como grupos genitivos,[4] como em *casa da minha irmã*, em oposição aos grupos qualificativos, que podem ser exemplificados por *casa de muitos cômodos*. Essas últimas construções simplesmente "qualificam", "especificam" ou "modificam" o núcleo nominal (MÜLLER, 1997, p.112), isto é, são adjuntos, mas que não admitem a paráfrase com *cujo* e nem a substituição pelo pronome possessivo.

[4] *Grupos genitivos* são aqui entendidos como aqueles que permitem os seguintes testes de identificação: (i) a relativização por *cujo*; (ii) a possibilidade de substituição por um pronome possessivo.

21.2.1 O sistema de indicação de posse no português brasileiro culto

A tradição gramatical apresenta o sistema de posse como um sistema bem simples e em equilíbrio, no qual os pronomes possessivos indicam, de forma inequívoca, o que cabe ou pertence a cada uma das pessoas gramaticais. Desse modo, em gramáticas pedagógicas (e.g. CUNHA; CINTRA, 1985), geralmente aparecem quadros de cuja leitura se pode presumir que existe uma correlação direta entre cada um dos pronomes pessoais e os possessivos, apesar de haver um sincretismo no que concerne ao singular e ao plural da 3ª pessoa, como ilustrado no Quadro 4.

PRONOME PESSOAL SUJEITO	PRONOME POSSESSIVO
eu	meu/minha/meus/minhas
tu	teu/tua/teus/tuas
ele/ela	seu/sua/seus/suas
nós	nosso/nossa/nossos/nossas
vós	vosso/vossa/vossos/vossas
eles/elas	seu/sua/seus/suas

Quadro 4 - O sistema de expressão de posse na norma padrão do português

Além disso, o fato de o pronome *vós* ter caído em desuso desde o século XVIII estendeu esse sincretismo para a 2ª pessoa do plural. Essa ambiguidade entre a 2ª e 3ª pessoas do discurso, no que concerne à indicação de posse, tornou-se crítica com a substituição generalizada do pronome de 2ª pessoa do singular *tu* pela forma pronominal *você*, proveniente da expressão nominal *vossa mercê* (CINTRA, 1972; FARACO, 1982). Essa substituição, predominante em grande parte do território brasileiro, no que concerne à norma culta, teria levado à situação expressa no Quadro 5.

PRONOME PESSOAL SUJEITO	PRONOME POSSESSIVO
eu	meu/minha/meus/minhas
você	*seu/sua/seus/suas*
ele/ela	*seu/sua/seus/suas*
nós	nosso/nossa/nossos/nossas
vocês	*seu/sua/seus/suas*
eles/elas	*seu/sua/seus/suas*

Quadro 5 - O sistema de expressão de posse do português com a queda dos pronomes pessoais *tu* e *vós*

Tais mudanças teriam conduzido a um sistema instável, na medida em que as formas dos pronomes possessivos não permitiriam a distinção entre a segunda e a terceira pessoas do discurso, ficando só a primeira pessoa com formas exclusivas de indicação de posse. Isso teria levado a uma crescente utilização da expressão genitiva *dele* (e formas flexionadas *dela/deles/delas*) para indicar a posse em relação à 3ª pessoa (SILVA, 1982, 1884, 1996; PERINI, 1985; CERQUEIRA, 1993, 1996). Tal substituição ocorreria para eliminar a ambiguidade, exemplificada pela frase apresentada em (14a.) abaixo, que seria, para um brasileiro, ambígua, já que não fica claro se o falante está se referindo à esposa de Paulo ou à do seu interlocutor. Portanto, a substituição de *seu* por *dele*, como em (14b.), tornaria clara a referência da posse para a 3ª pessoa.

(14) a. Eu falei com o Paulo e com a sua esposa.

b. Eu falei com o Paulo e com a esposa dele.

Essa motivação funcional tem sido destacada por vários autores, como Perini (1985, p.4), que defende o caráter funcional da introdução das formas analíticas com a preposição *de* no PB, sendo, em princípio, agramaticais as expressões analíticas de indicação de posse que não cumprissem essa função. Muller (1997, p.108-109) também destaca o fato de que a indicação de posse com o grupo *de* + pronome só é permitida com o pronome *ele* e com as "formas de tratamento" *você(s)* e *a gente*, já que essas últimas não exibem mesmo flexão de caso. Mas, nesses dois últimos casos, a falta de motivação funcional explicaria a sua baixa frequência na norma culta brasileira, de modo que, mesmo se referindo ao sujeito com as formas *você* e *a gente*, o falante culto brasileiro ainda usa preferencialmente os pronomes possessivos *seu* e *nosso*, respectivamente.

Isso se comprova na norma culta com quantificação da variação na referência da posse para a 2ª pessoa do singular que Neves (1993, p.164) apresenta, identificando três variantes (*de você*, *seu* e *teu*) relacionadas com o pronome sujeito *você*. Os seus resultados apontam para um largo predomínio da forma canônica *seu*, com 86,1% do total de ocorrências, contra 5,6% para a forma *teu* e 8,3% para a forma *de você*. No que concerne à 1ª pessoa do plural, ainda no universo da norma culta, as descrições conhecidas parecem indicar que o uso de *nosso* e flexões não tem sido inibido pela implementação da forma *a gente* nas outras funções sintáticas. Monteiro (1994, p.211) exclui a forma analítica de referência à posse de primeira pessoa gramatical, uma vez que, embora admita a existência da forma *da gente* na norma urbana culta, afirma que "*da gente* ainda não parece ameaçar a existência do pronome *nosso*". Opinião semelhante é sustentada por Omena (1996, p.191), que, analisando a inserção da forma *a gente* com várias funções sintáticas na fala de indivíduos com nível médio de escolaridade da cidade do Rio de Janeiro, conclui que, na função de adjunto adnominal, "ainda predomina o uso de *nosso*".

Neves (1993, p.165), realizando uma análise variacionista acerca do sistema de posse no PB, com dados do Projeto NURC, registra que a forma *nosso* e flexões ocorreu quarenta e duas vezes, ao passo que a forma *da gente* é representada apenas cinco vezes, demonstrando a correlação entre a forma inovadora na função de sujeito e a conservadora como genitivo: "*a gente* andava de bicicleta, era o esporte predileto *nosso*".

Com isso, parece que as pressões funcionais acabam por definir o sistema de expressão de posse na norma culta brasileira, já que nela só a expressão analítica *dele(a)(s)* ganhou vitalidade, exatamente como meio de desfazer o sincretismo na expressão de posse da 2ª e da 3ª pessoas. Além desse caso, tem-se destacado apenas o uso da forma analítica *de vocês*, que cumpriria a função de operar a distinção entre um e mais de um possuidor. As outras formas analíticas disponíveis, *de você* e *da gente*, não ganharam vitalidade no uso culto, já que a sua introdução não preencheria qualquer função no sistema. O Quadro 6 representa o sistema pronominal de expressão de posse em uso na norma culta brasileira.

PRONOME PESSOAL SUJEITO	PRONOME POSSESSIVO
eu	meu(s)/minha(s)
você	seu(s)/sua(s) ~ teu(s)/tua(s)
ele/ela	*dele(a)*
nós	nosso/nossa/nossos/nossas
vocês	seu(s)/sua(s) ~ *de vocês*
eles/elas	*deles(as)*

Quadro 6 - Sistema de expressão de posse em uso no português urbano culto

Porém, a observação das formas em uso nas variedades populares do PB para expressar a posse revela um quadro bem distinto desse, com um amplo emprego das formas analíticas *de você(s)* e *da gente*, ao lado de um residual uso de formas analíticas bastante divergentes com *de eu* e *de nós*, que estão totalmente ausentes do uso culto. Com isso, o português popular afasta-se de uma evolução interna da língua das elites escolarizadas e aproxima-se das línguas crioulas de base portuguesa da África, onde abundam expressões analíticas de pronomes pessoais regidos pela preposição *de* (*di*) — cf. seção 21.1 acima. Desse modo, a análise do processo mais amplo de variação no sistema pronominal de posse das variedades populares do PB só poderá ser levada a cabo se o seu algoritmo integrar o processo histórico do amplo e intenso contato entre línguas que marca a sua formação. A descrição e análise variacionista da expressão de posse no português afro-brasileiro apresentadas nas seções seguintes pretendem demonstrar isso.

21.2.2 O sistema de indicação de posse no português afro-brasileiro

Nesta seção, será feita uma descrição horizontal das formas de indicação de posse encontradas atualmente na fala das comunidades rurais afro-brasileiras, a partir de um levantamento exaustivo de todas as formas de expressar a posse referentes a todas as pessoas do discurso em uma amostra de fala compostas por 24 entrevistas sociolinguísticas com moradores de quatro comunidades rurais isoladas de diferentes regiões do Estado da Bahia: Helvécia, no Extremo Sul; Barra-Bananal, na Chapada Diamantina; Cinzento, no Sudoeste, próximo ao município de Vitória da Conquista; e Sapé, no Recôncavo Baiano, próximo à cidade de Valença.[5]

No que concerne à posse referente à 1ª pessoa do singular, o sistema de posse do português brasileiro urbano culto parece não admitir variação, pois não se registram usos como "casa *de eu*"; "casa *de mim*", exceto na fala das crianças. Entretanto, no *corpus* do português rural afro-brasileiro, foram identificadas ocorrências da forma analítica *de eu*. Em um universo de 943 ocorrências de expressões de posse com referência à 1ª pessoa do singular, foram encontradas três formas analíticas, apresentadas abaixo:

(15) Num aumentô, a menina aumentô, ela tá do tamaim *de eu*. (SP-05)[6]

(16) No caso *de eu*, todo dia eu saio pra trabaiá, plantá mandioca, colhê guaraná. (SP-06)

(17) Mas no caso *de eu*, dá pra encher o pé, por que a gente sobe lá pra cima. (SP-06)

Esse reduzido uso da forma analítica pode ser interpretado como um vestígio de um passado em que as formas analíticas de expressão de posse eram muito mais frequentes em comunidades de fala brasileiras constituídas em um contexto de contato linguístico amplo e maciço, atingindo todas as pessoas do discurso, mesmo aquelas cujas formas não estavam comprometidas por qualquer processo de sincretismo, o que elimina qualquer hipótese de evolução estrutural interna motivada por pressões funcionais.

A descrição e a análise da expressão de posse com relação às demais pessoas do discurso reforçam a ideia de uma substituição ampla de formas sintéticas por formas analíticas, como aquela que se observou em situações mais radicais de contato linguístico. No que diz respeito, por exemplo, à posse referente à 1ª pessoa do plural, foi encontrado no *corpus* analisado um alto índice de ocorrências da

---

[5] Uma descrição detalhada dessas comunidades é feita no capítulo 2 deste livro; a metodologia empregada na constituição das amostras de fala é apresentada no capítulo 5.

[6] Os exemplos são retirados da amostra de fala analisada, com a indicação do falante, através de um número, e da sua comunidade, através das siglas: HV, para Helvécia; RC, para Rio de Contas; CZ, para Cinzento; e SP, para Sapé.

forma analítica *da gente*, exibindo uma variação bem estruturada com a forma sintética *nosso* e flexões, que será analisada na próxima seção.

Já a análise da expressão da posse referente à segunda pessoa do discurso foi dificultada em função do tipo de entrevista que constitui as amostras de fala analisadas.[7] Desse modo, foram encontradas apenas 59 ocorrências com formas indicativas de posse referente à segunda pessoa, sendo a grande maioria com dados de discursos reportados, a exemplo de "o INPS preguntô: aonde o senhô qué ***seu*** pagamento?" (HV-20). As 59 ocorrências estão assim distribuídas: 04 com a forma *teu* e flexões, 05 com a forma *de vocês* e 50 com a forma *seu* e flexões.

Quanto à posse referente à terceira pessoa do discurso, seja no singular ou no plural, ocorreu um uso predominante da forma perifrástica *de + ele* (*dele* e flexões), com 322 ocorrências, contra 59 da forma sintética pronominal (*seu* e flexões); devendo-se destacar que essas últimas só foram registradas em contexto de posse genérica e indeterminada, como nos exemplos abaixo, nos quais o uso da forma analítica *dele* (e flexões) nesse contexto seria agramatical:

(18) Não... não, cada um tem ***seus*** tipo de comida. (CZ-03)

(19) Gosto, senhô, demais... ININT as coisa tá muito dificil. Hoje em dia, quem tem ***seu*** pedacinho de terra, se ficá dexano assim abandonado, num dá não... (HV-07)

(20) É, a enxadinha... enxadinha e depois de certos tempo pra cá que Nosso Senhô botô essa escolada no mundo, todo mundo hoje em dia sabe assiná ***seu*** nome...(RC-24)

Assim, o sistema de posse no português afro-brasileiro pode ser sistematizado como no Quadro 7.

PRONOME PESSOAL SUJEITO	PRONOME POSSESSIVO
eu	*meu(s)/minha(s)*
você	seu(s)/sua(s) ~ teu(s)/tua(s)
ele/ela	*dele(a)*[8]
nós	*nosso(a)(s) ~ da gente*
vocês	seu(s)/sua(s) ~ *de vocês*
eles/elas	*dele(s)/ dela(s)*

Quadro 7 - Sistema de expressão de posse no português afro-brasileiro

Como se pode ver, as formas analíticas predominam na 3ª pessoa, e estão em variação equilibrada com as formas sintéticas, na 1ª e 2ª pessoas do plural. Na

---

[7] Cf. capítulo 5 deste livro.

[8] O uso da forma pronominal ocorre em contextos muito específicos, tendo uso apenas em posses genéricas ou indeterminadas.

1ª pessoa do singular é que a forma sintética parece estar definitivamente consolidada, bem como na 2ª pessoa do singular com a alternância *seu/teu*. De qualquer forma, o registro de variantes bem divergentes como *de eu* e *de nós*, que ainda se encontram, não apenas no português afro-brasileiro, mas também no português popular das regiões mais afastadas do interior do Brasil, parece apontar para um passado em que, em consequência dos efeitos do contato entre línguas, as formas analíticas introduzidas pela preposição *de* com um pronome do caso reto concorriam com as formas sintéticas tradicionais em todas as pessoas do discurso. A análise variacionista da expressão de posse em relação à 1ª pessoa do plural — para a qual se encontraram as variantes *nosso(a)(s)*, *da gente* e *de nós* — apresentada nas próximas seções buscará reunir elementos que demonstrem, no plano histórico e linguístico, os efeitos do contato entre línguas na formação das variedades populares do português brasileiro.

21.3 Análise sócio-histórica da indicação de posse com referência à 1ª pessoa do plural no português afro-brasileiro

A amostra de fala utilizada para a análise variacionista da expressão de posse com referência à 1ª pessoa do plural foi constituída por 24 entrevistas sociolinguísticas, gravadas e armazenadas em meio digital, e estruturada segundo os seguintes fatores: faixa etária, sexo, comunidade, nível de escolaridade e estada fora da comunidade; sendo dois informantes por célula.[9] A variável dependente constitui um grupo ternário, assim exemplificado:

(25) É. ***Nossos*** pai é irmão do ôto. (CZ-11)

(26) A situação ***da gente*** vai melhorá mais... cada vez mais. (CZ-03);

(27) Verdadêra terra ***de nós*** é esse lá... aucalipto, não. (HV-13)

A distribuição dessas três variantes em termos de sua frequência na amostra estudada é apresentada na Tabela 1.

Tabela 1 - Frequência geral das variantes de expressão de posse com referência à 1ª pessoa do plural no português afro-brasileiro

FORMAS DA EXPRESSÃO DE POSSE					
nosso(a)(s)		*da gente*		*de nós*	
Nº ocorr.	%	Nº ocorr.	%	Nº ocorr.	%
48	44	57	52	05	04%

[9] Cf. capítulo 5 deste livro.

A escassez de dados com a forma *de nós* se explica em função do avançado processo de assimilação de padrões linguísticos e culturais urbanos que se acelerou na segunda metade do século XX e que estaria eliminando as principais marcas decorrentes de formação multilíngue das comunidades rurais afro-brasileiras (cf. LUCCHESI, 2001a e o capítulo 1 deste livro). Porém, a baixa frequência de uso da forma analítica *de nós* não é um empecilho para que se possa fazer ilações sobre o seu encaixamento no processo sócio-histórico de formação dessa variedade do PB. Nesse sentido, pode-se pensar que o uso da forma analítica *de* + pronome *nós* teria facilitado a implementação da forma analítica *da gente*. Isso poderia explicar por que o uso de *da gente*, que encontra resistência na norma culta (cf. seção 21.2.1. acima), seja a variante mais usada nas comunidades rurais afro-brasileiras, com uma frequência superior à metade do total de ocorrências.

Com base nesse raciocínio, partiu-se da hipótese de que a forma analítica *de nós*, mais distante do uso culto e mais próxima dos crioulos portugueses, seria a forma crioulizante e, portanto, mais antiga, devendo ser mais frequente na fala de informantes mais velhos; ao passo que as formas *nosso* e *da gente* são ambas inovadoras e mais frequentes entre os informantes mais jovens, de acordo com o seguinte esquema evolutivo:

de nós → *nosso* ~ *da gente*

Haveria, portanto, na história dessas comunidades, dois processos paralelos de mudança: um, com a aquisição da forma *nosso*; e, outro, com a aquisição de *da gente*; ambos motivados pela adoção de padrões externos às comunidades. Assim, a forma *de nós* foi sendo substituída pela forma sintética *nosso*, à proporção que aumentava o contato das comunidades antes isoladas com o mundo exterior e a influência do padrão linguístico urbano; esse processo foi facilitado pelo uso mais frequente da forma *nós* como sujeito na fala das pessoas mais velhas das comunidades; paralelamente, a forma *a gente*, em expansão nas variedades urbanas, também penetra nas comunidades na função de sujeito (cf. capítulo 19 deste livro), tendo o seu deslocamento para a função de adjunto adnominal (como na construção "terra *da gente*") facilitado pela existência anterior no sistema de formas analíticas de marcação de posse, como *de nós*.

A variável **faixa etária** é aquela que fornece os dados cruciais para o estabelecimento dos processos de mudança em progresso (LABOV, 1981; LUCCHESI, 2001b). Os dados dessa variável na amostra analisada são apresentados na Tabela 2.

Entre os falantes mais velhos, com idade média de aproximadamente 70 anos, predominam as formas analíticas, com a presença residual da variante mais divergente da norma culta, *de nós*. Adotando-se a "hipótese clássica" do tempo aparente (NARO, 2003, p.44), segundo a qual o falante fixaria o seu padrão de

comportamento linguístico entre os 15 e os 20 anos,[10] o comportamento linguístico desses falantes de 70 anos refletiria a situação da comunidade de fala entre 1945 e 1950. Dessa forma, até meados do século XX, ainda se podiam encontrar vestígios dos efeitos mais notáveis do contato entre línguas na fala dessas comunidades. A partir daí, a forma mais crioulizante *de nós* foi desaparecendo em favor das variantes *nosso* e *da gente*. O predomínio de *nosso* entre os falantes da faixa 2, de 42 a 57 anos, com 68% do total de ocorrências aponta para uma onda anterior de mudança com a introdução dessa variante, há aproximadamente 50 anos, já que esses informantes teriam fixado os padrões de sua gramática nos anos de 1960, sofrendo os primeiros efeitos do processo de urbanização verificado no Brasil, no período. A introdução da forma *da gente* seria posterior e associada ao processo de implementação de *a gente* como pronome pessoal, predominando entre os informantes da faixa 1, com idade entre 24 e 37 anos, na frequência de 72%.

Tabela 2 - Formas de expressão de posse segundo a variável faixa etária

FAIXA ETÁRIA	FORMAS DA EXPRESSÃO DE POSSE					
	nosso(a)(s)		*da gente*		*de nós*	
	Nº oc./ Total	%	Nº oc./ Total	%	Nº oc./ Total	%
de 20 a 40 anos	10/36	28%	26/36	**72%**	0/36	0%
de 40 a 60 anos	26/38	**68%**	12/38	32%	0/38	0%
de mais de 60 anos	12/36	33%	19/36	53%	5/36	**14%**
Total	48/110	44%	57/110	52%	5/110	4%

Por outro lado, os resultados da variável *sexo* forneceram outros elementos que reforçam o caráter conservador da variante *de nós*, como se pode depreender a partir da leitura da Tabela 3.

Tabela 3 - Formas de expressão de posse referente à 1ª pessoa do plural segundo a variável sexo

SEXO	FORMAS DE EXPRESSÃO DE POSSE					
	nosso(a) (s)		*da gente*		*de nós*	
	Nº oc./ Total	%	Nº oc./ Total	%	Nº oc./ Total	%
Feminino	22/48	46%	21/48	44%	5/48	**10%**
Masculino	36/72	**50%**	36/72	**50%**	0/72	-
Total	48/110	44%	57/110	52%	5/110	04%

Os dados mostram que não foi encontrada nenhuma ocorrência da forma crioulizante *de nós* na fala dos homens, ficando essa forma restrita à fala das

[10] Enquanto Naro (2003) situa o momento da fixação do padrão aos 15 anos, Chambers e Trudgill (1980, p.165) afirmam que "a validade do [tempo aparente] depende crucialmente da hipótese de que a fala das pessoas de 40 anos hoje reflete diretamente a fala das pessoas de 20 anos há 20 anos atrás".

mulheres. Esse fato ajusta-se à ideia de que as mulheres, nas comunidades rurais brasileiras, têm um comportamento mais conservador, por estarem mais restritas ao ambiente familiar e da agricultura doméstica, deslocando-se menos que os homens para os centros urbanos e tendo menos contato com o mundo exterior. Isso pode explicar por que os homens também superam as mulheres no uso de *da gente*, já que essa seria a variante que estaria ganhando mais espaço nos grandes centros urbanos. Os processos paralelos de mudanças que têm consolidado no português afro-brasileiro as formas concorrentes *nosso* e *da gente* foi objeto de uma análise variacionista, cujos resultados são apresentados na próxima seção.

21.4 *Nosso* e *da gente* no português afro-brasileiro: uma análise variacionista

Retirando-se da base de dados as ocorrências de *de nós*, focalizou-se, em uma nova análise probabilística, somente as variantes *nosso(a)(s) / da gente*, as duas de maior vitalidade na gramática atual das comunidades de fala analisadas. Nessa análise variacionista, foram propostas quinze variáveis explanatórias, das quais nove de natureza linguística e cinco sociais. O programa de regras variáveis VARBRUL selecionou apenas seis como relevantes do ponto de vista estatístico: (i) comunidade; (ii) paralelismo formal; (iii) tipo de posse; (iv) faixa etária; (v) estada fora da comunidade; (vi) referencialidade, aqui apresentadas na ordem de sua seleção pelo programa de cálculo probabilístico.

21.4.1 A comunidade

Os resultados da variável **comunidade** revelaram que a implementação da variante *da gente* nas comunidades rurais afro-brasileiras tem uma distribuição diatópica significativa, como se pode ver a partir dos números da Tabela 4.

Tabela 4 - Influência da variável comunidade no uso da variante *da gente* na expressão de posse referente à 1ª pessoa do plural no português afro-brasileiro

COMUNIDADES	Nº DE OC./TOTAL	FREQUÊNCIA	PESO RELATIVO
Cinzento	6/29	21%	.12
Rio de Contas	9/18	50%	.18
Helvécia	20/35	57%	.66
Sapé	22/23	96%	.94
Total	57/105	54%	.59

Nível de significância: .033.

Constata-se, nas comunidades de Cinzento e Rio de Contas, uma maior vitalidade da forma pronominal sintética *nosso(a)(s)*, que se reflete nos baixos pesos relativos associados à variante inovadora *da gente*, respectivamente, .12 e .18. Essas duas comunidades são as mais afastadas do litoral, o que reforça a hipótese de que a implementação da forma *da gente* teria ocorrido em paralelo com a implementação de *a gente* na função sintática de sujeito, cujos resultados indicam que o ponto irradiador tenha sido a região do Recôncavo, que, do século XVII ao XIX, concentrou um grande contingente de escravos africanos em suas plantações de cana-de-açúcar e de fumo, expandindo-se primeiramente pelo litoral, para depois avançar sobre o interior do Estado (cf. capítulo 19 deste livro). Assim, pode-se explicar por que a comunidade de Sapé é aquela que exibe a maior frequência de uso de *da gente*, com peso relativo de .94, seguida de Helvécia, com peso relativo de .66, pois Sapé está próxima do Recôncavo Baiano, e a comunidade de Helvécia está situada no Extremo Sul, próxima ao litoral.

21.4.2 O paralelismo formal

No escrutínio do encaixamento linguístico da variação entre *nosso* e *da gente*, a variável selecionada pelo VARBRUL em primeiro lugar como estatisticamente significativa remete ao conceito de *paralelismo formal*. O paralelismo formal foi proposto por Scherre e Naro (1993, p.2) para dar conta de "uma tendência de formas gramaticais semelhantes ocorrerem juntas (cf. Schiffrin, 1981)". Assim, partiu-se da premissa de que a forma selecionada para a expressão de posse estaria correlacionada à seleção da forma de referência pessoal mais próxima, de modo que, nos contextos em que o falante expressasse a referência pessoal por meio da forma *a gente*, expressaria também a posse por meio da forma *da gente*, do mesmo modo que, quando utilizasse o pronome *nós*, utilizaria a forma sintética *nosso(a)(s)*. Os resultados quantitativos obtidos favoreceram essa hipótese, como se pode ver na Tabela 5.

Tabela 5 - Distribuição das ocorrências *nosso(a)(s)* e *da gente* segundo a variável paralelismo formal (correlação com o pronome pessoal sujeito) no português afro-brasileiro

FORMA DO PRONOME PESSOAL MAIS PRÓXIMA	FORMA DE EXPRESSÃO DE POSSE					
	da gente			*nosso (a) (s)*		
	Nº oc./ Total	%	P.r.	Nº oc./ Total	%	P.r.
nós	7/32	22%	.05	25/32	**78%**	**.95**
a gente	38/50	**76%**	**.87**	12/50	24%	.13
Total	45/82		—	37/82	45%	—

Nível de significância: .033.

Constata-se que, de fato, a frequência de uso das variantes está condicionada à escolha da forma de referência pessoal, pois os resultados da frequência das formas praticamente se invertem a depender da forma de referência pessoal utilizada, isto é, *a gente* favorece o uso de *da gente*, na ordem de 76% das ocorrências, enquanto *nós*, o uso de *nosso(a)(s)*, com 78%. Esses resultados percentuais são confirmados pelos pesos relativos, que isolam o efeito de cada fator, considerando-se a interferência simultânea de todas as variáveis independentes. Desse modo, confirma-se a hipótese da estreita relação entre o sistema de referência pessoal e o de posse no processamento linguístico de cada sentença.

21.4.3 O tipo de posse

A segunda variável selecionada pelo VARBRUL foi estruturada a partir de uma análise de três tipos de posse classificados: inalienável, material e abstrata. A posse **inalienáve**l é aquela em que o possuído é inerente ao possuidor, não podendo ser adquirido ao longo do tempo. A **abstrata** é aquela cujo possuído é algo que não pode ser comprado, mas que pode ser adquirido pela experiência, podendo ser de caráter moral, afetivo, social, intelectual e espiritual. Já a **material** é aquela em que o possuído é algo que pode ser comprado a qualquer momento. Os exemplos seguintes ilustram respectivamente esses três tipos de posse:

(28) Não... sempre... a gente trabaiava com os pai ***da gente*** mais a mãe, né... (HV-12)

(29) Às vez, amiga ***da gente***, né, que trabalha em Texêra e tal, liga. (HV-01)

(30) É.. e... e por aqui... Livramento por aqui mesmo é... é... é o pasto ***da gente***. (RC-24)

cujos resultados quantitativos estão na Tabela 6.

Tabela 6 - Distribuição das ocorrências de *nosso(a)(s)* e *da gente* segundo a variável tipo de posse no português afro-brasileiro

TIPO DE POSSE	FORMAS DE EXPRESSÃO DE POSSE					
	da gente			*nosso (a) (s)*		
	Nº de oc./ Total	%	P.r.	Nº oc./ Total	%	P.r.
Inalienável	25/31	81%	**.88**	6/31	19%	.12
Material	13/30	43%	.51	17/30	57%	.49
Abstrata	19/44	43%	.19	25/44	57%	**.81**
Total	57/105	54%	—	48/105	46%	—

Nível de significância: .033.

Observando-se os resultados quantitativos da Tabela 6, constata-se que o contexto mais favorecedor da forma *da gente* é a expressão da posse inalienável/inerente com o peso relativo de .88. Isso pode ser explicado na medida em que nos contextos com posse inalienável, que são posses com o traço [+natural], seja frequente a expressão vernacular *da gente*, ao passo que nas abstratas, com o traço [–natural], o falante busque a expressão por meio de recursos mais formais, utilizando a forma pronominal sintética, *nosso(a)(s)*, que na expressão da posse abstrata obteve um peso relativo de .81.

21.4.4 A faixa etária

Os resultados da variável **faixa etária** trouxeram novos elementos sobre o processo de implantação das variantes *da gente* e *nosso* nas comunidades analisadas, como se pode ver nos resultados da Tabela 7.

Tabela 7 - Influência da variável faixa etária no uso da variante *da gente* na expressão de posse referente ao "eu-ampliado"

FAIXA ETÁRIA	Nº OC./ TOTAL	FREQUÊNCIA	P. R.
24 a 37 anos	26/36	72%	.66
42 a 57 anos	12/38	32%	.14
63 a 107 anos	19/31	61%	.81
Total	57/105	54%	—

Nível de significância: .033.

A forte presença da variante *da gente* entre os falantes com mais de 60 anos (com frequência de 61% e peso relativo de .81) indica que a introdução dessa forma analítica teria sido facilitada na comunidade em função do uso da forma mais antiga *de nós*, pois o sistema já contaria com uma forma analítica para expressar a posse em relação à 1ª pessoa do plural. Desse modo, pode-se postular que houve uma primeira onda de mudança em favor de *da gente*, que já estava em curso no início do século XX e se estendeu até a década de 1960. Por outro lado, o predomínio da forma *nosso* entre os falantes com idade entre 42 e 57 anos aponta para uma segunda onda de mudança em favor dessa variante, no período de 1960 até 1980. De lá pra cá, estaria em curso uma nova onda de implementação de *da gente*. Portanto, o quadro aponta para movimentos cíclicos de introdução das formas *nosso* e *da gente*, em substituição à antiga forma crioulizante *de nós*. A existência na comunidade de uma forma analítica teria facilitado a entrada da forma *da gente*, o que não aconteceu na norma culta brasileira, na qual o possessivo *nosso* tem se mantido em uso, apesar do avanço da forma nominal *a gente* sobre o pronome canônico *nós* na função de sujeito.

21.4.5 Estada fora da comunidade

Ao se destacar os informantes que viveram pelo menos seis meses fora da comunidade (geralmente trabalhando em grandes centros urbanos, ou no Sudeste do país), busca-se capturar o processo de difusão de uma mudança que advém de influências externas à comunidade. Assim, os falantes que viveram fora da comunidade liderariam o processo de mudança em favor da variante inovadora. Os resultados obtidos para a variante *da gente* com essa variável na rodada do VARBRUL estão na Tabela 8.

Tabela 8 - Influência da variável estada fora da comunidade no uso da variante *da gente* na expressão de posse referente ao "eu-ampliado"

ESTADA FORA DA COMUNIDADE	Nº OC./ TOTAL	FREQUÊNCIA	P. R
Sim	21/44	48%	.75
Não	36/61	59%	.31
Total	57/105	54%	—-

Nível de significância: .033.

Os resultados da Tabela 8, com peso relativo de .75 para o fator estada fora da comunidade a favorecer o uso de *da gente*, reforçam a ideia de que o uso dessa variante inovadora decorre, por um lado, da influência de padrões externos às comunidades rurais afro-brasileiras, relacionando-se ao incremento de *a gente* na função de sujeito. Por outro lado, a alta frequência de uso de *a gente* para expressar a posse no SP *da gente* estaria sendo favorecida pelo fato de a gramática original dessas comunidades já prever o uso de formas analíticas para expressar a posse, diferentemente do que ocorre nos dialetos urbanos, como já se expôs acima.

21.4.6 Nível de referencialidade

Em seu processo de gramaticalização, através do qual a expressão *a gente* passou a concorrer com o pronome *nós* como forma de expressão da 1ª pessoa do plural, essa forma inovadora conservou em grande parte de suas ocorrências o traço semântico [–específico]. Isso se deve ao fato de essa expressão nominal, em seu sentido original, referir-se a uma coletividade (cf. capítulo 19 deste livro). Consequentemente, é de se esperar que o emprego do SP *da gente*, para a expressão de posse com relação à 1ª pessoa do plural, também refletisse essa preferência para os contextos de referência genérica, mantendo-se o contexto de referência específica como um contexto de conservação da forma canônica *nosso*.

Para testar essa hipótese, a variável explanatória nível de referencialidade foi estruturada da seguinte maneira:

(i) referente com o traço [+genérico]

(31) É, ela estraga as vista ***da gente*** que bebê ela. (HV-13)

(ii) referente com os traços [–genérico]

(32) Acho que é mais barato que aqui na ***nossa*** cidade, aí... (CZ-03)

Os resultados quantitativos obtidos para o emprego de *da gente* considerando essa variável estão reunidos na Tabela 9.

Tabela 9 - Influência da variável nível de referencialidade no uso da variante *da gente* na expressão de posse referente à 1ª pessoa do plural

TRAÇOS SEMÂNTICOS	Nº OC./ TOTAL	FREQUÊNCIA	P. R
[+genérico]	12/16	75%	.92
[–genérico]	45/89	51%	.45
Total	57/105	54%	—-

Nível de significância: .033.

Como se pode ver, o uso da expressão analítica que tem em seu núcleo a expressão nominal *a gente* é bastante favorecido no contexto de referência genérica (com frequência de uso de 75% do total de ocorrências e peso relativo de .92), mantendo-se o contexto de referência específica e definida como de maior utilização da forma sintética *nosso(a)(s)*, com peso relativo de .55 para essa variante. Entretanto, deve-se salientar que, mesmo nesse contexto, as duas variantes exibem praticamente a mesma frequência de uso (51% do total de ocorrências para a variante sintética contra 49% para a variante analítica), o que demonstra que o emprego de *a gente* como adjunto adnominal que expressa a posse está bem mais adiantado no português afro-brasileiro do que em outras variedades do PB.

21.5 Considerações finais

Esta análise sociolinguística do sistema de posse na gramática das comunidades rurais afro-brasileiras isoladas revelou processos de variação e mudança que fornecem evidências empíricas consistentes para a polarização sociolinguística do Brasil e para a relevância dos processos de variação e mudança induzidos pelo contato entre línguas na formação histórica da realidade linguística brasileira.

Nesse sentido, é significativa a diferença entre os processos de variação e mudança que se observam hoje na norma culta e na norma popular do PB.[11] Na

[11] Para uma visão da polarização sociolinguística do PB entre uma norma culta e uma norma popular, veja-se Lucchesi (1998, 2001a, 2002a, 2002b, 2006).

norma culta, destaca-se a introdução da forma analítica *dele* (e formas flexionadas *dela/deles/delas*) para a expressão da posse com referência à terceira pessoa do discurso, em detrimento da forma sintética *seu* (e formas flexionadas *sua/seus/suas*). Essa substituição teria como fato desencadeador mais notável o emprego da forma *seu* com referência à segunda pessoa do discurso, decorrente da substituição do pronome pessoal do caso reto *tu* pelo pronome *você*. Portanto, a alteração mais significativa na pauta dos pronomes pessoais na norma culta brasileira decorreria de um processo estrutural de mudança em cadeia que poderia ser esquematizado da seguinte maneira: (i) substituição de *tu* por *você*; (ii) utilização preferencial na fala do possessivo *seu* com referência ao interlocutor, criando-se a ambiguidade na utilização de *seu* com referência à 3ª pessoa; (iii) utilização da forma analítica *dele* para evitar a ambiguidade. Portanto, o quadro de variação que se observa na norma culta resulta essencialmente de pressões funcionais decorrentes de mudanças no paradigma dos pronomes pessoais.

Já o quadro de variação e mudança que se observa na norma popular brasileira é bem distinto do que se observa na norma culta e aponta para um processo sócio-histórico igualmente distinto de formação dessa variedade do PB. Na fala dos segmentos populares, para além do uso generalizado da forma analítica *dele*, o emprego das formas *de você(s)* e *da gente*, diferentemente do que se observa na norma culta, é muito frequente. Além disso, encontra-se ainda o emprego residual das formas *de eu* e *de nós*, bastante desviantes do padrão da língua portuguesa e que apontam fortemente para processos de mudança desencadeados por fatores externos ao sistema linguístico, particularmente para o contato entre línguas, como ficou demonstrado com a descrição dos sistemas de expressão de posse dos crioulos de base lexical portuguesa da África.

Esse paralelo com os sistemas de posse dos crioulos portugueses da África revelou que o português afro-brasileiro se formou a partir de um processo de *transmissão linguística irregular* do tipo leve, e não a partir do processo radical em que se formam os crioulos típicos. Desse modo, não há registro no Brasil da solução radical em que uma única forma do pronome é empregada para todos os casos, inclusive o genitivo, em formas do tipo *vocês filhos* (com o sentido de 'seus filhos') ou *nós casa* (com o sentido de 'nossa casa'). O paralelo que se pode traçar entre a estratégia de expressar a posse por meio de um pronome pessoal regido pela preposição *de* presente em todos os crioulos de base lexical portuguesa aqui descritos e a presença significativa dessa estratégia de expressão de posse nos padrões de comportamento linguístico das comunidades rurais afro-brasileiras isoladas (e com menor intensidade em toda a norma popular brasileira), sendo usada inclusive para todas as pessoas do discurso, até para a 1ª pessoa do singular (um uso que pode ser considerado agramatical para um falante culto), torna a

hipótese da transmissão linguística irregular não só a mais factível, como também a mais simples, portanto aquela que prevalece nos esquemas mais aceitos de avaliação das explicações científicas.

Portanto, a presença residual das expressões *de eu* e *de nós* na fala das comunidades rurais afro-brasileiras remete para um estágio de formação dessas comunidades em que, em função do tipo de variação e mudança que ocorre nos processos de transmissão linguística irregular desencadeados pelo contato maciço entre línguas, predominaram as formas analíticas em detrimento das formas sintéticas com flexão de caso. Porém, já no início do século XX, as formas analíticas, como *de eu* e *de nós*, estariam se perdendo, com o avanço das formas sintéticas, como *meu* e *nosso*; e, no caso da 1ª pessoa do plural, a substituição de *nós* por *a gente* na função de sujeito, possibilitou a substituição de *de nós* por outra forma analítica: *da gente*. E a grande vitalidade desse SP para expressar a posse no português afro-brasileiro, diferentemente do que se observa na norma culta, em que a grande vitalidade de *a gente* na função de sujeito não encontra paralelo na função de adjunto adnominal para expressar a posse, pode ser explicada pelo fato de a gramática das comunidades afro-brasileiras já conter a possibilidade estrutural de expressão de posse por meio de um SP com a preposição *de*.

A análise sociolinguística da variação *de nós~da gente~nosso* reuniu elementos que iluminaram o processo de variação e mudança em que formas analíticas e sintéticas se alternam para expressar a posse na gramática das comunidades rurais afro-brasileiras isoladas. As conclusões deste capítulo acabam por confinar com as conclusões do capítulo anterior, bem como com vários outros capítulos da segunda parte deste livro em que se descreve a gramática do português afro-brasileiro, demonstrando que as abordagens que se encerram nos chamados fatores internos, baseadas exclusivamente em uma lógica funcional imanentista, não são capazes de dar conta dos processos de variação que se observam no português brasileiro, sobretudo se o foco da análise transcender a norma culta e incluir os padrões de fala da grande maioria do povo brasileiro: a chamada norma popular do PB. Para compreender satisfatoriamente o amplo e complexo cenário linguístico brasileiro, bem como a sua história, é preciso fazer o que se espera de qualquer sociolinguista: integrar os processos da estrutura linguística nas disposições mais amplas das condições reais de uso da língua e nos contextos sócio-históricos em que esse uso linguístico se formou.

Conclusão

Dante Lucchesi

As análises que compõem este livro buscaram individualizar os padrões de fala vernácula observados nas comunidades rurais afro-brasileiras isoladas como uma variedade específica no universo sociolinguístico do Brasil. A denominação **português afro-brasileiro** ganha sentido quando se identificam as características linguísticas e os traços sócio-históricos que particularizam tal variedade, mesmo no cenário já marcado da norma popular brasileira. Que características linguísticas destacam as comunidades afro-brasileiras isoladas do conjunto das comunidades de fala popular do interior do país? Que particularidades sócio-históricas possibilitaram a emergência de tais propriedades individualizadoras? Tais propriedades permitem vislumbrar processos pretéritos e localizados de crioulização do português no Brasil? Em que medida tais processos de variação e mudança induzidos pelo contato se espraiaram para o conjunto de variedades que vieram a conformar o **português brasileiro** como uma variedade linguística notavelmente distinta do **português europeu**? E atualmente, como se comportam essas comunidades de fala no bojo dos processos sociolinguísticos em curso no interior do país? Essas foram as principais questões abordadas neste livro.

A hipótese reitora foi a de que as variedades populares do português brasileiro, formadas em situações de multilinguismo, são marcadas por processos de variação e mudança derivados do processo de **transmissão linguística irregular** que ocorre com a socialização de uma variedade de segunda língua adquirida precariamente por falantes adultos em situações de sujeição e com a nativização desse modelo defectivo de segunda língua entre os descendentes desses grupos dominados. Dessa forma, as variedades populares do português brasileiro apresentariam características semelhantes às das línguas crioulas do Caribe e da costa ocidental da África, pois se teriam formado em contextos sócio-históricos semelhantes em função do colonialismo europeu e da escravidão africana. Contudo, as especificidades e sobretudo a complexidade da sociedade brasileira, na qual se destaca o alto grau de mestiçagem e a situação ambígua e contraditória dos mestiços nesse contexto histórico, não permitiram que a crioulização do português ocorresse de forma representativa e duradoura, pois não se reproduziram aqui as mesmas condições

nas quais vicejaram, por exemplo, as línguas crioulas do Caribe: sociedades com uma presença ínfima de colonizadores europeus, nas quais as populações africanas subjugadas eram fortemente segregadas. Condições semelhantes teriam ocorrido no Brasil ao longo do século XVII, período em que mais provavelmente teriam ocorrido processos de pidginização e crioulização do português, em contextos como o do Quilombo dos Palmares, desbaratado no ano 1695. Mas, até o início do século XVIII, a população do Brasil não passava de pouco mais de trezentos mil habitantes. E as profundas mudanças que já se iniciam nesse século com o ciclo da mineração viriam a alterar profundamente os cenários da colonização do Brasil, com a imigração maciça de colonos portugueses e o aprofundamento da miscigenação racial, tanto que, no século XVIII, a população do Brasil cresceu cerca de onze vezes, e, em meados do século XIX, praticamente a metade da população do país já era constituída por mestiços.

Mas, se as particularidades sócio-históricas do Brasil não propiciaram o surgimento de comunidades crioulófonas, não se pode argumentar de forma consistente contra o fato de que o português que se viabilizou socialmente como a língua dos afrodescendentes e dos indiodescendentes mestiços e endógamos não tenha sofrido alterações significativas em função dos processos de transmissão linguística irregular desencadeados pelo contato entre línguas. Esse português profundamente alterado seria a língua de cerca de dois terços da população brasileira no final do século XIX, uma população iletrada e concentrada no interior do país, pois, até o início do século XX, mais de oitenta por cento da população do Brasil vivia no campo. Nas pequenas cidades costeiras, concentrava-se a elite branca e letrada, que buscava reproduzir aqui os modelos linguísticos da antiga Metrópole, como parte de um projeto de dominação excludente e racista, fundado na violenta exploração do trabalho que sobreviveu à abolição da escravatura e se estende até os dias atuais, em que ainda se surpreendem condições de trabalho análogas à escravidão em grandes propriedades rurais no interior do país. Delineia-se, assim, o quadro histórico da polarização sociolinguística do Brasil. De um lado, uma reduzida elite ciosa em reproduzir aqui os hábitos linguísticos de Portugal. De outro, o português profundamente alterado pelo contato entre línguas falado pela maioria da população.

Decerto que a elite branca não ficou imune aos efeitos do contato. A relação com escravos/empregados domésticos (dentre os quais figuram as famosas amas de leite), o convívio dos filhos dos senhores com os filhos dos escravos/empregados, entre outras formas de socialização, acabaram por contaminar a norma culta brasileira, aprofundando as suas diferenças com a sua congênere europeia. A vinda de mais de três milhões de imigrantes europeus e asiáticos entre o final do século XIX e o início do século XX também criou importantes vetores de difusão de estruturas populares provenientes do contato linguístico para a norma culta

brasileira (LUCCHESI, 2001a). Por outro lado, o grau de alteração do português falado por afrodescendentes e por indiodescendentes também teria variado bastante consoante cada contexto histórico específico. Certamente, as alterações mais profundas ocorreram em torno das *plantations* e nos quilombos, situados nas regiões mais afastadas e remotas do interior, e foram menos intensas nos meios urbanos e mesmo no universo da mineração (CASTRO, 1990). Portanto, configura-se na virada dos 1900 um quadro sociolinguístico extremamente complexo e altamente matizado, de variedades que foram afetadas em maior ou menor grau pelo contato entre línguas, delineado-se um *continuum* desde as comunidades de fala de origem quilombola até o português lusitanizado das elites urbanas, passando pelo português das comunidades rurais mestiças, das vilas do interior, da elite rural, das classes urbanas subalternas, etc.

A industrialização e a urbanização da sociedade brasileira, impulsionadas após o fim da República Velha, em 1930, alteraram bastante o cenário sociolinguístico brasileiro. Hoje, mais de oitenta por cento dos cerca de 180 milhões de brasileiros vivem nas cidades do país, em sua grande maioria, nas condições degradantes de suas periferias urbanas. O grosso dessa população é proveniente do êxodo rural, criando a figura do indivíduo **rurbano**, com fortes implicações linguísticas (BORTONI-RICARDO, 2005). Com isso, o que era variação diatópica, na dicotomia cidade/campo, tem-se transformado em variação diastrática, ou, mais precisamente, tem o seu perímetro reduzido à oposição centro/periferia, nas grandes cidades.

O rurbano é um ser híbrido, no qual se encontram formas linguísticas rurais mais diretamente afetadas pelo contato e as formas do padrão urbano culto que ele vem assimilando em seu processo de urbanização. A influência dos padrões linguísticos urbanos tem crescido enormemente, superando em muito o perímetro das regiões metropolitanas e atingindo todas as áreas do país, através da poderosa ação dos meios de comunicação de massa, que já constituem certamente o quarto poder da república. O deslocamento populacional, com as migrações sazonais de indivíduos que saem do interior para trabalhar nas grandes cidades e depois retornam à sua terra natal, também contribui significativamente com a difusão dos padrões linguísticos urbanos, assim como a expansão da rede pública de ensino, em que pesem as suas precárias condições de funcionamento no interior do país. Em suma, assiste-se, desde meados do século XX, a um processo de nivelamento linguístico do país que, em última instância, resulta da integração econômica e cultural de novos segmentos sociais, ou seja, da ampliação do universo da cidadania. Trata-se de um conjunto de **mudanças de cima para baixo**[1] nas quais as variantes urbanas mais valorizadas socialmente começam a ser assimiladas pelos segmentos

[1] Para o conceito de *mudança de cima para baixo*, veja-se Labov (2008 [1972]).

das classes mais baixas, consoante o seu grau de inclusão social. Concomitantemente, vai-se abandonando o uso das variantes mais divergentes derivadas do contato entre línguas, que normalmente são marcadas socialmente com uma avaliação explicitamente negativa.

Por outro lado, a dinamização dos círculos sociais da elite brasileira em função da industrialização e da urbanização do país arrefeceu a influência linguística lusitana, e movimentos literários como o Modernismo reagiram explicitamente ao purismo gramatical. Porém, os segmentos dominantes mais conservadores e seus veículos de comunicação de massa mais empedernidamente reacionários continuam a alimentar uma espécie de **racismo linguístico**, como parte de seu projeto de exclusão social, superexploração do trabalho e concentração de renda. Dessa contradição entre uma distensão dos usos linguísticos na chamada **norma culta** e a manutenção pela tradição gramatical (ortodoxa ou midiática) de um padrão linguístico normativo adventício e anacrônico (a **norma padrão** brasileira de matiz lusitano) resulta um sentimento de **insegurança linguística** que se abate sobre toda a sociedade (em particular, os seus segmentos letrados) e que se traduz na máxima recorrente de que "o brasileiro não sabe falar o português". Portanto, está em curso um tênue processo de convergência dos dois polos da realidade sociolinguística brasileira (a **norma culta** e a **norma popular**) ainda permeado por profundas contradições, e cuja consecução passa pela ampliação dos direitos sociais e por uma efetiva distribuição das riquezas do país. Em outras palavras, a **democratização linguística** do Brasil é indissociável da real **democratização** da sociedade brasileira.

Assumir a **polarização sociolinguística** do Brasil tem fortes implicações para a pesquisa acerca da formação histórica da realidade linguística brasileira.[2] Como só foi afetada indiretamente pelo contato entre línguas, a norma culta brasileira exibe atualmente processos de variação e mudança que são fortemente condicionados por fatores estruturais, funcionais e discursivos. Porém, algumas análises contidas neste livro demonstraram que as formas mais divergentes da fala popular não se enquadram nesses esquemas explicativos estruturais e funcionais, pois resultam de processos mais radicais de mudança que extrapolam os limites da evolução interna da língua. Apesar disso, a grande maioria das análises diacrônicas sobre o **português brasileiro** focalizam apenas a história linguística dos segmentos ditos cultos, concentrando-se nesses fatores internos. Para isso, conjugam-se a forte tradição formalista e imanentista ainda hegemônica na pesquisa linguística e o fato de que os registros históricos existentes são quase que exclusivamente da

---

[2] Essa visão da polarização sociolinguística do Brasil vem sendo sistematizada por Lucchesi há mais de quinze anos (1994, 1996, 1998, 2001a, 2002a, 2002b, 2006, 2007) e tem sido encampada por outros linguistas, como Mattos e Silva (2004).

chamada língua culta. Desse modo, o conjunto de análises reunidas neste livro abre uma importante frente para a pesquisa sobre a história linguística do Brasil. Por um lado, desviam o foco da língua da elite letrada para as variedades linguísticas usadas pela grande maioria da população; variedades estas historicamente ignoradas ou desqualificadas como formas corrompidas, degradadas ou degeneradas da língua nacional. Por outro lado, resgatam a importância dos chamados fatores externos para a compreensão da história da língua, pois não se pode compreender a formação das variedades do português brasileiro, sem colocar no centro da cena as mudanças desencadeadas pelo contato linguístico massivo e radical.

Como se disse anteriormente, até o início do século XX, pelo menos oitenta por cento da população do Brasil vivia no interior, sendo constituída majoritariamente por índios, africanos e seus descendentes mestiços e endógamos. Nesse universo, o letramento era mínimo, e a normatização linguística, no geral, inexistente. Para Guy (1981a), não se trata, frente a esse quadro, de argumentar como o português se teria crioulizado, mas como ele teria deixado de se crioulizar. Os capítulos iniciais deste livro enfrentaram essa questão, argumentando no sentido de que não houve um processo representativo e duradouro de crioulização do português, porém o antecedente histórico do português popular contemporâneo, que Mattos e Silva (2004) chamou de **português geral brasileiro**, teria passado por processos significativos de mudança desencadeados pela situação de "multilinguismo generalizado" em que se formou, sem contudo crioulizar-se plenamente, caracterizando o que tem sido denominado aqui **transmissão linguística irregular de tipo leve**.[3]

A diferença fundamental entre um processo de transmissão linguística irregular de tipo leve e um processo radical que dá ensejo à formação de uma língua crioula é que, enquanto neste último a reestruturação gramatical é de tal monta que produz uma nova variedade linguística qualitativamente distinta da língua-alvo (da qual recebe o grosso do seu léxico),[4] o resultado, no primeiro caso, é a formação de uma variedade da língua-alvo afetada por mudanças induzidas pelo contato. Nesse sentido, este livro também oferece uma significativa contribuição para o campo da teoria acerca do contato entre línguas, definindo um protótipo dessa variedade linguística afetada pelo contato entre línguas que não chega a ser uma língua crioula típica.

Como bem observou Roberts (1997), uma língua crioula se define menos por suas propriedades estruturais do que por sua história. O processo em que se formam as línguas crioulas, aqui denominado **transmissão linguística irregular**,

---

[3] De uma forma muito apropriada, Silva Neto (1951, 1963) recorre ao conceito de semicrioulo para definir esse antecedente histórico do português popular brasileiro.

[4] Cf. Rougé (2008).

caracteriza-se essencialmente pela erosão gramatical de uma língua adquirida precariamente como segunda língua por um agrupamento de falantes adultos, seguida da recomposição gramatical dessa língua que se torna a língua nativa dessa comunidade que se forma na situação de contato. No caso dos crioulos mais radicais, o acesso aos modelos da língua-alvo é muito restrito, de modo que muito pouco da estrutura gramatical dessa língua-alvo passa para a nova variedade linguística. Assim, toda a reestruturação gramatical que dará forma à língua crioula é proveniente das estruturas adotadas pelos falantes adultos com base na sua competência linguística nativa ou da ação da faculdade da linguagem durante o processo de nativização desse modelo defectivo de segunda língua entre as crianças que nascem na nova comunidade de fala. Formadas assim, as línguas crioulas compartilham um conjunto de soluções para uma série de mecanismos gramaticais, tais como: sistema de tempo modo e aspecto (com partículas pré-verbais), mecanismos de regência e complementização oracional (com a serialização verbal), indicação da categoria nominal de número (com pluralizadores pré-nominais), etc.[5] Ou seja, na falta de acesso aos dispositivos gramaticais da língua-alvo, a recomposição é feita com a gramaticalização de itens do reduzido léxico disponível. Por outro lado, as línguas crioulas têm em comum a ausência de certos dispositivos mais formais e abstratos da gramática, tais como: concordância verbal, concordância nominal e flexão de caso. Essa simplicidade morfológica sugere que a gramática das línguas crioulas contém apenas os dispositivos e as propriedades essenciais ao funcionamento de uma língua natural (BICKERTON, 1999). Dessa forma, as línguas crioulas também compartilhariam a opção por parâmetros não marcados da gramática universal: ordem fixa SVO, ausência de movimento, realização obrigatória do sujeito pronominal, ausência de cliticização dos complementos verbais, etc. (ROBERTS, 1997). Portanto, as línguas crioulas podem ser caracterizadas em termos estruturais da seguinte maneira: ausência de morfologia sem valor referencial, parâmetros sintáticos não marcados e dispositivos gramaticais resultantes de processos recentes de gramaticalização (LUCCHESI, 2003b, 2008c).

A grande diferença entre o processo de transmissão linguística irregular de tipo leve e a crioulização deriva de uma diferença no nível de acesso aos modelos da língua-alvo. A transmissão linguística irregular de tipo leve caracteriza-se por um maior acesso aos modelos da língua-alvo, não se verificando a violenta situação de segregação e isolamento necessária ao surgimento de uma comunidade de fala crioulófona. Havendo uma maior interação entre dominantes e dominados, o nível de erosão gramatical da língua-alvo é menor. Além disso, mesmo que haja uma forte erosão da língua-alvo no momento inicial de sua aquisição como segunda

[5] Como argumentado por Roberts (1997) e outros, tais estruturas não são, entretanto, privativas das línguas crioulas.

língua pelos adultos dos grupos dominados, os dispositivos gramaticais da língua-alvo tendem a ser reincorporados em função do maior acesso potencial a esses dispositivos por parte das crianças que nascem na situação de contato. Esse esquema se ajusta perfeitamente ao fato histórico reconhecido de que os filhos dos escravos nascidos no Brasil se integravam muito mais na sociedade brasileira do que seus antepassados africanos. Nesse contexto, praticamente não ocorrem processos de gramaticalização dos dispositivos originais que caracterizam as línguas crioulas. E mesmo os mecanismos gramaticais mais abstratos e formais não desaparecem totalmente, ocorrendo antes um quadro de ampla variação em seu uso. Dessa forma, não chegam a ocorrer mudanças nos parâmetros sintáticos da língua-alvo em favor de parâmetros não marcados, constatando-se mais uma vez um quadro de profunda variação.

Uma das maiores contribuições deste livro é fornecer, através do conjunto de análises aqui reunidas, uma formalização analítica empiricamente fundamentada do processo de transmissão linguística irregular de tipo leve. As análises dos dezesseis aspectos da morfossintaxe da gramática das quatro comunidades rurais afro-brasileiras aqui reunidas revelaram que, nos casos de transmissão linguística irregular de tipo leve, a única porção da estrutura gramatical que é significativamente afetada é aquela constituída por mecanismos gramaticais mais abstratos e formais que não têm valor referencial, o que será denominado aqui **gramática aparente**. E, mesmo nesses casos, não ocorreu a eliminação total desses mecanismos, não se criando as condições para que se efetivassem alterações nos parâmetros sintáticos da língua-alvo. E praticamente não se observaram casos de formação de dispositivos originais na gramática. Para além da contribuição teórica para as pesquisas acerca do contato entre línguas, essa formalização analítica fornece uma chave interpretativa crucial para a compreensão da formação histórica dos padrões linguísticos atuais da maioria da população brasileira.

Dedutivamente, é possível chegar-se à conclusão de que houve processos de crioulização do português na história do Brasil, assim como é possível deduzir-se que houve ou há vida inteligente fora da Terra. Mas, da mesma forma como as absurdas escalas de tempo e espaço do universo não permitem a comprovação empírica da segunda assertiva, a ausência de registros históricos dos processos sociolinguísticos ocorridos em engenhos e quilombos do Brasil dos séculos XVII e XVIII não permite confirmar ou refutar cabalmente a primeira. Isso não significa que a história linguística dos afrodescendentes no Brasil não seja possível. Nesse ponto, este livro encontra o seu significado maior, pois as comunidades afro-brasileiras podem ser vistas hoje como verdadeiros sítios arqueológicos que fornecem valiosas indicações acerca do que se passou na maior parte do interior do Brasil entre os séculos XVIII e XX, período em que o povoamento do país ganhou um

verdadeiro impulso. E não deixa de refletir também o que se passou entre africanos, negros e mestiços nos pequenos núcleos urbanos no mesmo período.

Formadas por descendentes de escravos africanos que se mantiveram em torno de antigas plantações, ou que se agruparam em quilombos em locais recônditos e inacessíveis, muitas dessas comunidades mantiveram um contato reduzido com o mundo exterior até meados do século XX. Dessa forma, as comunidades rurais afro-brasileiras isoladas seriam as comunidades de fala que teriam passado por processos de transmissão linguística irregular mais intensos no conjunto atual das variedades linguísticas brasileiras. Mas, se assim podem ser vistas, isso não significa obviamente que as comunidades aqui analisadas sejam as que mais tenham sido afetadas pelo contato no universo estimado de aproximadamente oitocentas comunidades desse tipo que ainda se conservam no Brasil, fora as comunidades que já se desarticularam. Mesmo no universo reduzido de quatro comunidades, constataram-se diferenças nos efeitos do contato entre línguas identificados entre uma comunidade e outra. Além disso, não se pode perder de vista que as amostras de fala vernácula foram recolhidas nessas quatro comunidades, na década de 1990 e nos primeiros anos do século XXI. E as informações disponíveis sobre a comunidade de Helvécia indicam que as mudanças descrioulizantes podem ter ocorrido de forma bastante acelerada, mesmo em face do relativo isolamento dessas comunidades. Diante disso, é preciso distinguir as situações de transmissão linguística irregular do tipo leve das situações de *continuum post*-crioulo, e a comparação da situação das comunidades afro-brasileiras com o que ocorre, por exemplo, na Jamaica e na Guiana é, nesse sentido, bastante ilustrativa.

Nesses países em que houve um processo pleno de crioulização, observa-se, na situação atual de *descrioulização*, um *continuum* desde as variedades basiletais, que exibem as estruturas mais tipicamente crioulas, até as variedades acroletais, que se confundem com qualquer variedade não padrão do inglês (RICKFORD, 1987). Portanto, não se encontra no Brasil uma situação típica de descrioulização, com um *continuum post*-crioulo típico, pois, mesmo nas comunidades afro-brasileiras que supostamente passaram por um processo mais radical de reestruturação gramatical, não se encontram estruturas tipicamente crioulas, como partículas pré-verbais de tempo, modo e aspecto, verbos seriais e pluralizadores pré-nominais, que ainda se encontram nas variedades basiletais da Guiana e da Jamaica. Esse contraste constitui uma forte evidência contra a hipótese de Guy (1981a) de que o português se teria crioulizado no Brasil nos séculos XVII e XVIII, passando por um processo de descrioulização a partir do século XIX. Por outro lado, favorece a visão em que este livro se baseia.

A visão de que o que ocorreu no Brasil foi um generalizado processo de transmissão linguística irregular de tipo leve, pontilhado por processos de crioulização localizados e efêmeros, aumenta a importância do estudo das

comunidades afro-brasileiras em função de suas particularidades sócio-históricas. O fato de serem constituídas quase que exclusivamente por descendentes diretos de escravos africanos é garantia de que tenham passado por um processo de transmissão linguística irregular. O isolamento em que se conservaram é igualmente crucial, no sentido de os seus padrões de fala atuais ainda exibirem reflexos desse processo anterior decorrente das situações de multilinguismo de que derivam, mesmo que o processo de transmissão linguística irregular por que passaram tenha sido de tipo leve, sem ter atingido um nível de crioulização (o que parece ser o caso de todas, com a exceção de Helvécia). E isso não compromete o valor empírico dos dados recolhidos, pois, ao exibirem tais características, essas comunidades acabam por refletir processos muito mais representativos, que em maior ou menor grau permearam a grande maioria dos contextos sociolinguísticos do interior do país até os finais do século XX. Não apenas dos inúmeros quilombos e grandes engenhos e plantações de fumo e algodão, mas também os contextos das minas, das pequenas plantações cujos pequenos proprietários tinham em média de três a cinco escravos, das pequenas vilas, e até mesmo o universo dos africanos, crioulos mestiços do mundo subalterno das cidades.

Os resultados das análises sobre a morfossintaxe do português afro-brasileiro atual constituem uma importante contribuição para a compreensão da história das variedades populares do português brasileiro, não apenas daquelas que ainda hoje se conservam no interior do país, como também daquelas que se concentram nas periferias das grandes cidades, em função do êxodo rural ocorrido nos últimos cinquenta anos. E, considerando essa importância das comunidades rurais afro-brasileiras isoladas, as análises que integraram a segunda parte deste livro buscaram individualizar a gramática dessas comunidades no universo sociolinguístico brasileiro, tanto no plano linguístico, quanto no plano sociolinguístico.

No plano linguístico, buscou-se identificar os processos de variação que estariam mais diretamente ligados às situações de contato entre línguas. Nesse plano, uma estratégia das análises foi a de traçar paralelos com processos de mudança que caracterizam as línguas crioulas de base portuguesa, sobretudo da África, em virtude das semelhanças históricas que aproximam tais línguas do português afro-brasileiro. Por outro lado, foi feito também o cotejo com outras variedades do português brasileiro, para deslindar as diferenças quantitativas e qualitativas que constituem as balizas para a compreensão da formação histórica do universo sociolinguístico do Brasil, na medida em que pudessem indicar como as mudanças induzidas pelo contato se espraiaram, ou não, pelos diversos segmentos sociais. Assim, observaram-se diferenças qualitativas tanto em face da norma culta brasileira, quanto em face das demais variedades populares urbanas e rurais do Brasil; ou seja, observaram-se em algumas comunidades rurais afro-brasileiras isoladas processos de variação, como os que afetam a concordância de gênero no interior

do Sintagma Nominal (SN) e a concordância verbal junto à 1ª pessoa do singular, que estão ausentes, não apenas da norma urbana culta, como também da grande maioria de variedades populares urbanas e rurais, particularizando as comunidades rurais afro-brasileiras isoladas como uma variedade linguística do Brasil em função desses processos específicos de variação diretamente relacionados com o contato entre línguas.

Por outro lado, em alguns dos processos de variação induzidos pelo contato que se difundiram por todas as variedades linguísticas brasileiras, afetando inclusive a norma culta (como no caso da variação na concordância verbal e na concordância nominal de número), as análises buscaram identificar diferenças quantitativas que refletissem a maior intensidade com que as comunidades rurais afro-brasileiras foram afetadas pelo contato *vis-à-vis* às demais variedades cultas e populares do português brasileiro. Por fim, essas análises tiveram de enfrentar o desafio teórico de formalizar as maneiras como o contato afeta a estrutura da língua, através de um jogo dialético em que a análise se funda na teoria disponível no campo da crioulística para guiar a observação dos fatos, ao tempo em que as evidências empíricas obtidas possibilitam o avanço da reflexão teórica. Nessa frente de investigação, nem sempre foi possível estabelecer os nexos teóricos entre o contato e os fatos observados na estrutura da gramática das comunidades analisadas. Nesse caso, a contribuição desses trabalhos vem pela negativa, definindo aqueles aspectos da estrutura que, até então, parecem mais imunes aos efeitos do contato linguístico.

Uma outra forma de estabelecer uma relação histórica empiricamente motivada entre os processos de variação e mudança que se observam hoje nas comunidades rurais afro-brasileiras isoladas e as situações de contato linguístico massivo, abrupto e radical assenta na descrição do perfil sociolinguístico atual desses processos de variação e mudança. Os resultados das variáveis sociais na análise sociolinguística de vários aspectos da gramática das comunidades de fala analisadas revelaram um quadro de mudança em progresso no qual as formas mais características do contato linguístico estão sendo substituídas pelas formas do padrão urbano culto em função do processo de nivelamento linguístico que se processa por todo o interior do país pela influência dos grandes centros urbanos. Tal padrão de mudança desautoriza qualquer visão de uma deriva secular que fosse gradualmente erodindo as marcas flexionais da língua ao longo de inúmeras gerações. Ao contrário, o que se vislumbra é um processo histórico bem definido, no qual o contato linguístico produziu uma forte erosão gramatical até o século XIX e, a partir daí, com crescente intensidade, principalmente a partir de meados do século XX, foi se implementando a recomposição de algumas marcas flexionais e mecanismos gramaticais, em função da integração dessas comunidades ao sistema socioeconômico do país.

Portanto, as análises contidas neste livro reuniram uma massa significativa de dados e reflexões analíticas que permitem:

(i) identificar os processos de variação e mudança que hoje afetam as variedades do português brasileiro que foram induzidos por processos pretéritos de contato entre línguas;

(ii) individualizar as comunidades afro-brasileiras isoladas como uma variedade específica do português brasileiro particularmente afetada por esses processos de contato;

(iii) definir parâmetros para balizar as diversas variedades no universo sociolinguístico brasileiro, em função de diferenças quantitativas e qualitativas entre os processos de variação e mudança observados;

(iv) traçar um perfil sociolinguístico das comunidades rurais afro-brasileiras isoladas no contexto sociolinguístico dos processos de variação e mudança que atualmente afetam as variedades populares do português no interior do país.

Assim sendo, se buscará, nesta **conclusão**, fazer uma síntese dos resultados das análises linguísticas reunidas neste livro, bem como traçar o perfil sociolinguístico das comunidades analisadas que emerge dessas análises.

Caracterização linguística do português afro-brasileiro

Os resultados das análises linguísticas confirmaram de forma muito consistente o ponto de vista defendido ao longo deste livro. Mesmo nas comunidades de fala mais afetadas pelo contato entre línguas, confirmou-se o que vale para o conjunto das variedades populares do português brasileiro: os aspectos da estrutura linguística mais afetados pelo contato concentram-se na **gramática aparente**: a morfologia flexional que serve para marcar processos gramaticais mais abstratos e sem valor referencial, tais como a relação sujeito-verbo, as fronteiras sintagmáticas e as funções sintáticas dos constituintes. Nesse conjunto, destacam-se a variação na concordância de gênero e a variação na concordância verbal junto à 1ª pessoa do singular como processos de variação exclusivos de algumas comunidades rurais afro-brasileiras isoladas, como reflexo de um processo mais profundo de alteração promovido pelo contato linguístico. Na concordância nominal de número e na concordância verbal junto às demais pessoas do discurso, observam-se diferenças quantitativas significativas entre o português afro-brasileiro e as demais variedades rurais e urbanas do português popular brasileiro. Já em relação à norma culta, pode-se dizer que as diferenças quantitativas são de monta a configurar algumas

diferenciações qualitativas, que confirmam, por um lado, a atual polarização sociolinguística do Brasil e, por outro, os processos históricos diferenciados que conduziram à formação da norma culta e da norma popular brasileira, tendo a primeira sido afetada de forma indireta pelo contato do português com as línguas indígenas e africanas.

A polarização sociolinguística do Brasil também se manifesta na variação na flexão de caso dos pronomes pessoais. Nesse aspecto da gramática, enquanto a variação na norma culta se apresenta como o resultado de pressões estruturais e mudanças discursivas, a amplitude dos processos na norma popular brasileira em geral, e nas comunidades afro-brasileiras isoladas em particular, extrapolam os limites de um processo de mudança estrutural e discursivo, apontando para um processo mais radical de mudança, induzido pelo contato entre línguas. Confirma-se aí também a existência de dois vetores históricos independentes e diferenciados na formação da realidade sociolinguística brasileira: aquele que conduz à norma culta brasileira contemporânea e aquele em que se formaram as variedades populares do português brasileiro, não obstante as interações existentes entre os dois segmentos ao longo da história até os dias de hoje.

A variação na concordância verbal com a 1ª pessoa do singular aponta para um quadro pretérito de erosão de todo o paradigma da flexão de pessoa e número do verbo que teria caracterizado as comunidades rurais afro-brasileiras até o século XIX. Tal quadro encontra uma correspondência perfeita com todas as línguas crioulas de base lexical portuguesa da África e da Ásia, nas quais esse mecanismo gramatical foi totalmente eliminado. Em realidade, a perda da flexão verbal de número e pessoa é uma característica virtualmente universal, não apenas das situações de contato massivo entre línguas, mas de todas as situações de aquisição defectiva de segunda língua por parte de falantes adultos.

Diante de um paralelo tão notável, fica difícil argumentar de forma consistente em favor de uma deriva secular ou de vagas propensões estruturais. Tal visão torna-se ainda menos defensável quando se atualiza em um processo em que, como argumentam Naro e Scherre (1993, 2007) reproduzindo o velho esquema da Filologia do século XIX, mudanças fonéticas de perda de segmentos finais desencadeariam um processo de erosão morfológica que se difundiria, por analogia, para todo o paradigma flexional. Com efeito, é muito pouco plausível que uma desnasalização rarefeita pudesse ter produzido tamanho estrago em um paradigma flexional com um conjunto tão diferenciado de marcas flexionais. Para além de sua fragilidade teórica, tal esquema também está em flagrante conflito com as evidências empíricas disponíveis. O padrão sociolinguístico consistente de mudança em curso no sentido da recomposição das marcas flexionais que se tem observado em todas as variedades populares do português brasileiro nas últimas décadas nega claramente um processo gradual de perda das marcas flexionais ao longo dos

séculos. Além disso, essa visão de uma perda gradual por analogia não faz o menor sentido no contexto sociolinguístico do Brasil entre os séculos XVI e XIX, no qual uma grande massa de índios e africanos adquiriam o português precariamente como segunda língua. É evidente que o emprego das marcas flexionais por esses falantes do português como segunda língua em situações tão adversas seria praticamente nulo, sendo progressivamente recomposto pelos seus descendentes, em função do contexto histórico de maior inserção dos crioulos na sociedade brasileira *vis-à-vis* a seus antepassados africanos, como já se destacou acima.

A erosão das marcas flexionais de pessoa e número do verbo, associada à realização do pronome sujeito, é inexorável no contexto de contato linguístico massivo, tanto que tais marcas desaparecem (como aconteceu nas línguas crioulas da África e da Ásia), se, nas condições sócio-históricas em que essas comunidades de fala se formam, o acesso aos modelos gramaticais da língua do superstrato se mantém reduzido. Mas, como argumentado ao longo deste livro, os contextos históricos brasileiros, no geral, não atingiram níveis tais de segregação que propiciassem a eliminação completa da flexão de pessoa e número do verbo (não descartando a possibilidade de isso acontecer em algumas comunidades de fala isoladas, sobretudo entre os séculos XVII e XVIII). O quadro mais geral seria o de uma ampla variação atingindo todo o paradigma. Nesse cenário brasileiro, as marcas da 1ª pessoa do singular foram as mais resistentes e/ou as que mais rapidamente foram recompostas. Em um caso ou em outro, o fato é que a variação na concordância verbal com a 1ª pessoa do singular só foi encontrada, no universo das quatro comunidades aqui analisadas, na comunidade de Helvécia, a única para a qual há maiores evidências de um nível mais radical de reestruturação de tipo crioulizante. Esse cenário suscita uma questão teórica que se coloca no plano do que a teoria sociolinguística tem chamado de *problema das restrições* (WEINREICH; LABOV; HERZOG, 2006 [1968]): o fato de as marcas flexionais da 1ª pessoa do singular terem sido menos afetadas pelo contato entre línguas se deve a alguma restrição de ordem estrutural ou discursiva, ou constitui apenas uma particularidade do caso brasileiro? O cotejo com o processo histórico de erosão das marcas flexionais do verbo em inglês, em que a forma superveniente foi o *-s* da 3ª pessoa do singular, favorece a visão de que, em processos de erosão do paradigma flexional do verbo, as marcas de pessoa e número que serão perdidas ou conservadas resultarão mais de um acidente histórico (isto é, do valor social que cada marca pode eventualmente assumir, ou de sua consistência fonética particular, etc.) do que de condicionamentos universais da gramática ou do discurso.

No caso específico do português popular brasileiro, as marcas da 1ª pessoa do singular exibem uma grande vitalidade num quadro de violenta erosão do paradigma flexional de pessoa e número. O morfema *-s* da 2ª pessoa do singular desapareceu, mesmo em áreas, como o interior do Estado da Bahia, em que o

pronome sujeito *tu* se conserva em uso.[6] Isso confirma mais uma vez a evolução diferenciada e independente das normas culta e popular no Brasil, pois, na primeira, a perda desse morfema decorre da substituição do pronome *tu* pela forma de tratamento sincopada *você*. Essa polarização também se manifesta no uso do morfema da 1ª pessoa do plural, pois, se a falta de concordância do tipo *nós trabalha* está praticamente ausente da fala de indivíduos urbanos com nível superior de escolaridade, ela atinge um percentual de mais de oitenta por cento no português afro-brasileiro. Diferença expressiva também se encontra no cotejo das frequências do morfema de plural *-m*, empregado com sujeitos da 3ª pessoa do plural, nos dois extremos do universo sociolinguístico brasileiro. Entre os informantes do Projeto NURC do Rio de Janeiro, a frequência de aplicação da regra de concordância verbal é da ordem de 94%, enquanto no português afro-brasileiro é de apenas 16%. Entre esses dois extremos, existe um *continuum* de variedades populares rurais, rurbanas e urbanas e de falantes urbanos com nível médio de escolaridade, tecendo a complexa rede dos estratos sociolinguísticos do Brasil.

Pode-se argumentar ainda que essa enorme diferença quantitativa reflete uma diferença qualitativa entre os dois processos, o que significa dizer que se configurariam dois processos distintos de variação. Uma evidência a favor dessa hipótese é o fato de que na norma culta a variação está praticamente confinada aos contextos de posposição do sujeito junto aos verbos inacusativos, em ocorrências do tipo: *Só apareceu três candidatos*. Nesses contextos, a variação pode atingir o patamar de 35% mesmo entre os falantes urbanos com curso superior completo (GRACIOSA, 1991). Na maciça variação observada nas comunidades rurais afro-brasileiras isoladas, esse fator estrutural não desempenhou um papel relevante como inibidor da aplicação da regra, sendo mesmo definido como um fator neutro nos resultados do Programa das Regras Variáveis (peso relativo de .51).

Portanto, no que concerne à morfologia verbal de pessoa e número, as comunidades rurais afro-brasileiras isoladas destacam-se por serem a variedade linguística brasileira que exibe a maior erosão do paradigma flexional. Diferentemente do que ocorre na maioria das variedades populares brasileiras, a variação atinge o uso de todos os morfemas, inclusive o da 1ª pessoa do singular. Quantitativamente, o português afro-brasileiro seria a variedade com as menores frequências de uso da regra de concordância verbal, em decorrência de uma maior erosão ocorrida na formação dessas comunidades de fala em função da transmissão linguística irregular desencadeada pelo contato entre línguas. O quadro de profunda variação daí proveniente teria chegado aos dias atuais em função da situação de isolamento dessas comunidades até a segunda metade do século XX. De qualquer

[6] Mesmo no Sul do país, onde as marcas da 2ª pessoa do singular se conservaram, observa-se atualmente um acelerado processo de mudança no sentido do seu desaparecimento.

forma, deve-se destacar a diferença entre o quadro de variação decorrente de um processo de transmissão linguística irregular de tipo leve e a eliminação total da morfologia flexional de pessoa e número, que caracteriza as línguas crioulas, formadas em um processo mais radical de transmissão linguística irregular. A diferença repousaria essencialmente no grau de acesso aos modelos da língua-alvo. Não obstante o seu isolamento, o acesso aos modelos de português nas comunidades afro-brasileiras isoladas desde a sua formação, e sobretudo ao longo do século XX, sempre foi maior do que, por exemplo, o das comunidades crioulófonas de Cabo Verde e de São Tomé e Príncipe.

Ainda no plano da gramática aparente, outros paralelos significativos se verificam na estrutura do SN. O português afro-brasileiro exibe, provavelmente com maior intensidade, o mesmo padrão de marcação do plural que se observa na norma popular brasileira como um todo, no qual a marca recai quase que exclusivamente no determinante pré-nuclear. Esse é o mesmo padrão encontrado no crioulo português de Cabo Verde e no português dos tongas em São Tomé. Os tongas constituem um dos últimos episódios de sequestro de populações para o trabalho forçado promovido pelo colonialismo europeu. Trata-se da importação de trabalhadores braçais do continente africano para as plantações de café e cacau da Ilha de São Tomé, no final do século XIX e início do XX. Um potencial processo de pidginização/crioulização do português nesse agrupamento foi sustado pela melhoria de suas condições de vida em função dos avanços sociais alcançados ao longo do século XX. De qualquer modo, essa identidade quanto à forma de marcar o plural no SN com variedades que passaram inegavelmente por um processo de reestruturação gramatical decorrente do contato entre línguas enfraquece mais uma vez o argumento de um desenvolvimento estrutural autônomo, apenas acelerado pelo contato.[7] Por outro lado, deve-se observar também que a variação na concordância nominal de número se difundiu para as normas urbanas culta e semiculta, configurando um caso de variação estruturada no português brasileiro, que não se observa no português europeu, para o qual não há qualquer registro de variação estruturada em qualquer uma de suas variedades dialetais, diferentemente do que sugerem Naro e Scherre (2007).

Já a variação na marcação do gênero no SN destacaria algumas comunidades rurais afro-brasileiras isoladas no universo dos falares populares do interior do país. Mesmo no universo de observação desta pesquisa, só a comunidade de Helvécia, tida como a mais crioulizante, exibiu uma variação estruturada na marcação do gênero no SN, com ocorrências do tipo: *eu trabalha n**o** roça*. E, mais uma vez, o

[7] Ainda mais quando se afirma (NARO; SCHERRE, 1993, 2007) que essa evolução interna assenta sobre uma indefensável tendência na România Ocidental para a perda do *-s* final, pois essa região latinizada da Europa caracteriza-se exatamente por conservar o *-s* do acusativo plural latino!

paralelo com o que se observa com os casos típicos de crioulização foi muito significativo. Na crioulização do português no continente africano, a morfologia do gênero foi totalmente eliminada. A reestruturação gramatical das línguas crioulas gerou uma marcação lexical para a referência ao sexo dos seres animados, do tipo: *filho macho / filho fêmea*. No *continuum post*-crioulo, essa marcação lexical passou a sofrer a concorrência da flexão no núcleo nominal: *filho / filha*, não ocorrendo normalmente a flexão de gênero entre os determinantes e modificadores do nome: *kel fija bunitu* 'aquela filha bonita'.[8] A variação observada em Helvécia obedece ao mesmo padrão, com a marcação categórica do gênero nos núcleos nominais, e marcação variável nos determinantes e sobretudo nos modificadores pós-nominais. Tal paralelo enfraquece mais uma vez a ideia de uma tendência evolutiva interna. Aliás, no que concerne à categoria de gênero, a ideia de uma deriva secular de perda das marcas morfológicas é absolutamente insustentável diante da evidente expansão da morfologia de gênero desde o português arcaico até os dias atuais, com palavras outrora invariáveis, como *senhor*, *espanhol* e *português*, que passaram a se flexionar em gênero: *senhora*, *espanhola* e *portuguesa*.

Já no plano da concordância nominal em estruturas passivas e de predicativo do sujeito, observou-se igualmente uma variação mais ampla na marcação da categoria de número do que na categoria de gênero. E o nível de variação nessas estruturas, no geral, também foi mais alto do que o observado no interior do SN, o que pode ser explicado pela baixa coesão estrutural entre o termo especificador e o especificado: o sujeito, por um lado, e o predicativo ou particípio da passiva, por outro; conquanto tenha se observado o efeito do princípio da coesão estrutural, na medida em que o falante tende a fazer mais a concordância dos particípios passivos e predicativos quando aplica a regra de concordância no SN sujeito, do que quando não o faz. O cotejo entre variedades linguísticas brasileiras revelou que, se a variação na concordância de número em estruturas passivas e de predicativo do sujeito é alta na norma urbana semiculta (em torno de 50%), o uso dessa regra está praticamente ausente nas comunidades rurais afro-brasileiras isoladas, ratificando mais uma vez a ideia da polarização sociolinguística e de uma maior erosão dos elementos da gramática aparente nas variedades linguísticas brasileiras que se formaram em situações de contato entre línguas. Já em relação à concordância de gênero, a frequência de aplicação da regra nas comunidades afro-brasileiras é bastante elevada, sendo, ao que tudo indica, categórica nas normas urbanas culta e semiculta.

Essa assimetria no nível de variação entre as categorias de número e gênero remete novamente ao problema das restrições, com a seguinte questão: nos casos de erosão da morfologia das regras de concordância no interior do SN, a morfologia

---

8 Exemplo do crioulo de Cabo Verde.

do número é geralmente mais afetada do que a do gênero, ou isso é uma particularidade do caso brasileiro? Lucchesi (2000) procura explicar isso pelo fato de, em português, a marcação do número ser mais redundante e gramaticalizada do que a do gênero. Enquanto a imensa maioria dos nomes e adjetivos se flexiona em número, a flexão de gênero nessas classes de palavras é muito mais restrita. Por outro lado, defende-se que o gênero seria um traço mais inerente ao nome do que o número. Ou seja, há uma série de parâmetros a serem considerados para se estabelecer um padrão estrutural dos processos de erosão da morfologia flexional do SN em situações de contato entre línguas, fazendo dessas situações um excelente espaço para o diálogo entre a teoria da variação e a teoria da gramática, confirmando a reflexão teórica desenvolvida no quarto capítulo deste livro.

A flexão de caso dos pronomes pessoais, incluindo o caso genitivo expresso pelos pronomes possessivos, foi outro aspecto da chamada gramática aparente afetado pelo contato entre línguas, pois a flexão de caso não tem valor referencial, servindo apenas para marcar a função sintática do constituinte. No caso das línguas crioulas, a função sintática é definida geralmente pela ordem dos constituintes, com um enrijecimento do padrão SVO. Desse modo, é comum nas línguas crioulas que a mesma forma do pronome desempenhe tanto a função de sujeito, quanto as de complemento verbal e adjunto adverbial, bem como a função de genitivo, indicando a posse. Nas comunidades estudadas, o reflexo mais notável do contato entre línguas foi identificado em ocorrências do tipo: *depende* ***de eu***, *matei* ***tu***, *terra* ***de nós***. Tais formas, hoje pouco frequentes, podem ser vistas como remanescentes de um estágio anterior em que a variação na flexão de caso teria sido muito mais geral, sobretudo no processo de socialização do português entre os escravos africanos e de nativização dessa variedade de segunda língua entre os descendentes desses escravos.

Para além de sua relação histórica com as situações de contato entre línguas, essas formas também definem uma clivagem entre as variedades populares do português brasileiro e suas normas urbanas culta e semiculta. Nessas últimas, a redução na flexão de caso restringe-se aos reflexos da substituição dos pronomes *tu* e *nós*, respectivamente, pelas formas *você* e *a gente*, em função de essas últimas não exibirem flexão de caso, por conta de sua origem nominal. Para além desses casos, observa-se apenas a perda do clítico acusativo de 3ª pessoa (*o/a/os/as*) e o crescente uso de *ele* (*eles/ela/elas*) não apenas como objeto direto, mas também como dativo: *deu isso* ***para ela***. Essa mudança pode ser derivada de pressões estruturais e desencadeada por alterações na prosódia (CÂMARA JR., 1972). Da mesma forma que o único possessivo ameaçado na norma culta é *seu* (*seus/sua/suas*), crescentemente substituído pela expressão analítica *dele* (*dela/deles/delas*) na modalidade oral para desfazer a ambiguidade, em função da assunção de *você(s)* como pronome de 2ª pessoa.

A variação na flexão de caso dos pronomes pessoais no português afro-brasileiro (bem como na norma popular brasileira) não se restringe ao pronome da 3ª pessoa, afetando formas genuinamente pronominais como *eu*, *tu* e *nós* (e.g., *não gosta de* ***eu****, vai mais* ***tu****, viu* ***nós***), fugindo assim ao algoritmo do jogo cambiante dos fatores estruturais em que se pode encerrar a variação com o pronome de 3ª pessoa. Do mesmo modo, formas analíticas de expressão de posse como *de nós*, *de tu* ou mesmo *de eu*, que ainda ocorrem na fala popular, não podem ser explicadas pelo esquema funcional de evitar a ambiguidade. Portanto, se os processos de variação e mudança identificados na norma culta podem ser explicados por meio de alterações discursivas e prosódicas ou pressões estruturais e funcionais, as estruturas radicalmente divergentes encontradas no português afro-brasileiro não se enquadram em tais esquemas, demandando um processo de mudança mais profundo, como aqueles de que derivam as línguas crioulas, não obstante a grande diferença que se observa entre a eliminação quase total da flexão de caso dos pronomes que ocorre na crioulização típica e a variação residual que se observa mesmo nas comunidades afro-brasileiras isoladas.

No que concerne à gramática aparente da língua, pode-se dizer, com base no que se observa no caso brasileiro, que os mecanismos da flexão nominal e verbal são mais afetados em processos de transmissão linguística irregular do tipo leve do que a flexão de caso dos pronomes. O caso dos crioulos de base portuguesa também parece corroborar essa hipótese, já que os crioulos, como o de Cabo Verde, ainda conservam algumas poucas formas pronominais flexionadas em caso, ao passo que eliminaram totalmente a concordância nominal e verbal em suas variedades basiletais.

Se, no plano da gramática aparente, os efeitos do contato entre línguas nas variedades populares do português brasileiro são muito claros, definindo nitidamente as fronteiras sociolinguísticas do país, que separam as variedades urbanas cultas das variedades populares e rurais, identificar os efeitos do contato em outros níveis da estrutura gramatical tem se configurado um objetivo muito mais difícil. Nesse caso, configuram-se questões que ainda desafiam toda a teoria do campo. Não há ainda um consenso na crioulística sobre como o contato entre línguas afeta particularmente as orações relativas, ou o parâmetro do sujeito nulo. Esses temas envolvem questões muito amplas e gerais que ainda permanecem em aberto nos diversos campos da ciência da linguagem, tais como: as estruturas nucleares da gramática universal; a existência de estruturas (ou parâmetros) marcadas e não marcadas; ou mesmo a existência de processos universais na aquisição de segunda língua em situações adversas, entre outras. A rede complexa de interações que caracteriza tais processos nem sempre permite deslindar os efeitos dos inúmeros fatores envolvidos, e os fatos não se apresentam de uma forma nítida; mesmo assim, alguns resultados reunidos neste livro foram significativos.

No caso da realização do sujeito pronominal de referência definida, não se encontrou uma polarização nítida no cotejo entre a frequência de sujeito nulo nas comunidades rurais afro-brasileiras e na norma urbana culta, frustrando a expectativa inicial de que o índice de sujeitos nulos nas comunidades afro-brasileiras isoladas fosse muito inferior ao encontrado na norma culta, já que, ao que tudo indica, as línguas crioulas se caracterizam pela ausência de sujeitos nulos referenciais, em função da ausência de flexão verbal de pessoa e número.[9] No caso brasileiro, parece ter havido uma convergência entre o polo culto e o popular em função da substituição generalizada de *tu* por *você* e do emprego crescente de *a gente* na modalidade oral da língua. Dessa forma, a grande diferença entre a norma culta e a popular se restringe ao uso do morfema *-m*, que, a rigor, seria menos um morfema de pessoa do que um morfema de número.[10] Como resultado disso, não se observou uma diferença muito significativa entre as comunidades rurais afro-brasileiras isoladas e a fala urbana culta, na frequência de uso do sujeito nulo. Uma pesquisa coordenada que empregue uma mesma base metodológica na análise desse aspecto da gramática nas diversas variedades linguísticas brasileiras, bem como futuros avanços teóricos do campo, poderão conduzir a um quadro mais nítido sobre a questão.

Já o estudo sobre processos de relativização revelou que a estratégia vernacular das relativas com antecedente é a de uso do relativizador universal *que*. Este fato é interpretado como o fator principal para a ausência de relativas *pied-piping*, pois estas só são possíveis quando a relativização envolve um pronome relativo. Contudo, a preferência pela *relativa cortadora* não pode ser vista como uma característica particular das comunidades afro-brasileiras derivada da situação de contato, pois vários estudos têm mostrado ser este o tipo mais recorrente de relativa em diferentes e diversos dialetos sociais brasileiros. Além disso, não se pode perder de vista os resultados dos estudos de Keenan e Comrie (1977, 1979), que mostraram uma ampla variação, nas línguas humanas, no que diz respeito aos usos das estratégias *pied-piping* e *cortadora*, que podem estar presentes ou ausentes em línguas de ramos e famílias (in)dependentes. Os dados analisados também não forneceram evidências diretas para a hipótese de uma influência direta do contato entre línguas, até porque não se definiu ainda que estratégias específicas de relativização são características das línguas crioulas. O estudo de Alexandre e

---

9 Como informado no capítulo 6, que trata desse tema, diferenças metodológicas entre as abordagens podem ter comprometido a comparação.

10 Observa-se também uma grande diferença entre a norma culta e a norma popular no emprego do morfema da 1ª pessoa do plural *-mos*. Porém, essa grande diferença fica circunscrita ao reduzido uso de *nós* na função de sujeito. Desse modo, o cada vez mais geral emprego de *a gente* está nivelando as duas normas na redução do uso de um morfema específico para a 1ª pessoa do plural.

Hagemeijer (2001) sobre relativização em diferentes crioulos atlânticos de base portuguesa aponta para a presença das três estratégias citadas acima, com distribuição diferente entre os diversos crioulos. Porém, a ausência sistemática de relativas *pied-piping* na fala rural de afrodescendentes, aliada ao fato de relativas *pied-piping* fazerem parte das diversas variedades vernaculares do português europeu e de variedades cultas do português, não pode deixar de ser relevante para a questão da aquisição do português por africanos e afrodescendentes. Indo nessa direção, o que caracterizou a gramática dos falantes das comunidades analisadas foi a reanálise na qual os relativizadores com características pronominais passaram a ser formas privativas das relativas sem antecedente; enquanto o complementizador *que* se generalizou para as formas de relativas com antecedente. Essa reanálise se ajusta perfeitamente ao cenário de transmissão linguística irregular em que essas comunidades se formaram.

No geral, o estudo das construções pseudoclivadas e clivadas mostra que esses tipos de construção são atestados em línguas com sócio-histórias diferentes, não se apresentando, portanto, como um ponto estrutural relevante para a discussão dos efeitos do contato linguístico nas comunidades analisadas. O dado mais revelador foi o uso de clivadas sem cópula (*João que fez isto*), construção não atestada nas variedades diacrônicas e sincrônicas do português europeu. Esta é certamente uma inovação do português brasileiro, em relação à língua-alvo; no entanto, não é restrita dialetalmente no Brasil, sendo atestada em qualquer variedade falada brasileira. De qualquer forma, é o tipo de mudança que se enquadra na característica essencial dos processos de transmissão linguística irregular: a perda de partículas gramaticais sem valor referencial.

Também foi atestado o uso de vários tipos de tópico, construções que se inserem na interface do discurso com a sintaxe, estando, portanto, subordinadas ao contexto discursivo em que ocorrem. A tipologia apresentada sobre as diferentes estratégias de realização dos tópicos em nada difere, estruturalmente, do que se tem observado sobre o português brasileiro em geral. O fato de haver uma ampla realização de tópicos por pessoas mais velhas das comunidades indica que essas construções não são inovadoras nas comunidades em estudo. Nesse sentido, pode-se concluir que as situações de contato não afetam o uso de estruturas de topicalização. Ao contrário, pode-se pensar que as favorecem, na medida em que a linguagem emergencial das situações de contato apoia-se muito mais no modo pragmático do que no sintático.

O estudo sobre a realização anafórica do objeto direto de 3ª pessoa (categoria vazia *versus* pronome lexical) partiu do pressuposto de que o uso das construções de objeto nulo se deve ao contato linguístico presente na formação das comunidades rurais afro-brasileiras isoladas. De fato, o equilíbrio no percentual de frequência dessa variante em todas as faixas etárias (73% na faixa 1, 70% na faixa 2 e 74% na

faixa 3) mostra que é a variante preferida pelos falantes, além de indicar ser a variante mais antiga nesse dialeto.

Já as estratégias de negação sentencial usadas nessas comunidades são fortemente condicionadas por fatores linguísticos e discursivos. De um modo geral, o uso da negação dupla e da negação pós-verbal atende, prioritariamente, a uma função de negação de pressuposições; a negativa pré-verbal, considerada discursivamente neutra, é favorecida pelos contextos de *não resposta*, por *orações encaixadas* e pela *realização lexical dos argumentos verbais*. Contudo, quando os dados são testados com informantes de outras variedades do português brasileiro, observam-se as mesmas restrições.

No plano dos mecanismos de estruturação da sentença, as análises apresentadas neste livro tiveram de enfrentar um terreno ainda bastante inexplorado no campo das investigações sobre os efeitos gramaticais do contato entre línguas. Nesse sentido, o maior obstáculo foi a ausência na teoria do campo de formulações sobre como o contato afeta certos aspectos estruturais da língua-alvo, para além daquelas características que são mais proeminentes na crioulização: o sistema de tempo, modo e aspecto, os verbos seriais, os pluralizadores nominais, etc. Mas, em que pesem as dificuldades decorrentes da ausência de sistematização teórica acerca de vários dos aspectos aqui abordados, algumas formulações interessantes foram alcançadas, tais como: a correlação entre a falta de morfologia flexional de pessoa e número do verbo e a realização do sujeito pronominal; a possibilidade de perda da natureza pronominal do relativizador em situações de contato; e o incremento da topicalização e focalização em tais situações.

Mesmo definindo a formação das comunidades rurais afro-brasileiras isoladas como um processo de transmissão linguística irregular de tipo leve, isso não significa que esse processo não possa afetar mecanismos gramaticais com algum valor informacional e que se deva descartar completamente processos de gramaticalização e alteração paramétrica. Nesse sentido, identificou-se como exemplo do primeiro tipo, a variação no emprego da morfologia de subjuntivo. A distância de mais de sessenta pontos percentuais que separa a frequência de emprego das formas do subjuntivo na fala de indivíduos escolarizados das cidades do Sul do país da frequência de emprego dessas formas nas quatro comunidades rurais afro-brasileiras comprovam a um só tempo a polarização sociolinguística do Brasil e os graus diferenciados com que o contato entre línguas se insere na formação histórica das diferentes variedades do português brasileiro, demonstrando que a erosão gramatical desencadeada pela transmissão linguística irregular de tipo leve pode afetar também de forma significativa os morfemas com algum valor referencial.

Entretanto, o processo de variação aqui identificado que revelou os níveis de maior radicalidade dos processos desencadeados pelo contato entre línguas foi a identificação da alternância dativa nas comunidades de fala aqui analisadas,

mantendo produtiva a construção de objeto duplo (COD), que gera frases do tipo: *dei os meninos o remédio*. O paralelo claro com as línguas crioulas, nas quais a COD predomina, associado ao seu caráter *sui generis*, não apenas na história da língua portuguesa, mas de todas as línguas românicas, nas quais está ausente, não deixa margem à duvida de que a emergência da COD no português afro-brasileiro só pode ser o resultado do massivo contato do português com as línguas africanas. Além disso, o fato de tais construções serem agramaticais para um falante culto das grandes cidades brasileiras coloca a polarização sociolinguística do Brasil em grau inédito de radicalidade, pois pode-se postular uma variação paramétrica separando as normas culta e popular no Brasil. Este livro coloca entre os seus achados mais significativos a identificação desse processo de variação e mudança do português popular brasileiro associado ao contato entre línguas que configura um caso claro de reestruturação original da gramática, do que resulta uma alteração paramétrica. Mais do que qualquer outro aspecto da gramática, a alternância dativa liga decisivamente o português popular brasileiro aos crioulos portugueses da África.

Pode-se sumarizar os efeitos do contato entre línguas no quadro linguístico atual das comunidades rurais afro-brasileiras isoladas da seguinte maneira:

(i) ampla variação no uso da morfologia flexional sem valor referencial que afeta particularmente os mecanismos de concordância nominal de número e concordância verbal de pessoa e número, mas que se manifesta também na flexão de caso dos pronomes pessoais e na concordância nominal de gênero, de forma residual;

(ii) variação mais localizada nos morfemas que exibem um valor informacional mais restrito, que afeta especificamente a morfologia do subjuntivo;

(iii) efeitos difusos em processos de variação que afetam a realização do sujeito pronominal, as orações relativas e clivadas, e as estruturas de tópico e de negação;

(iv) casos raros de reestruturação gramatical e de alteração paramétrica circunscritos até o momento à identificação da alternância dativa, com um grau significativo de funcionalidade na gramática dos falantes mais velhos das comunidades estudadas.

Por outro lado, a distribuição social que define o encaixamento social desse processo de variação e mudança, não apenas nas comunidades de fala aqui estudadas, mas sobretudo no plano mais geral do conjunto das variedades diastráticas do português brasileiro, oferece indícios significativos de como a ideologia dominante opera sobre os usos linguísticos para legitimar o caráter excludente e profundamente desequilibrado da sociedade brasileira.

Caracterização sociolinguística do português afro-brasileiro

Os resultados de algumas das análises sociolinguísticas reunidas neste livro são de grande relevância para a compreensão da dimensão sócio-histórica da realidade linguística brasileira. Destaque-se, em primeiro lugar, o fato de que os aspectos morfossintáticos mais afetados pelo contato entre línguas são exatamente aqueles que estão na base das principais clivagens sociolinguísticas do Brasil atualmente: a erosão da concordância nominal e verbal. O segundo fato está intimamente ligado ao primeiro: é exatamente sobre esses aspectos que a avaliação social da variação linguística é mais nítida, com um julgamento explicitamente negativo dessas características da fala popular provenientes do contato entre línguas. Os achados contidos neste livro revelam, portanto, de forma inequívoca que o preconceito linguístico das elites brasileiras não tem apenas um caráter social excludente e discricionário, ele tem um caráter claramente racista, na medida em que condena ao estigma social precisamente as marcas linguísticas mais notáveis do caráter pluriétnico da sociedade brasileira. Pode-se falar, assim, de um **racismo linguístico** ainda vigente e que atua impunemente como instrumento de dominação ideológica na sociedade brasileira. Uma síntese da distribuição social das variáveis linguísticas e do perfil sociolinguístico das comunidades analisadas poderá demonstrar isso.

Não se observaram diferenças sociais relevantes nos processos de variação que ocorrem nas estruturas da gramática que não foram afetadas diretamente pelo processo de transmissão linguística irregular. As frequências de uso de sujeito nulo, de relativas cortadoras, de dupla negação e de estruturas de tópico e foco encontradas na norma culta não diferem muito das encontradas na norma popular e no português afro-brasileiro, não havendo a discriminação explícita das variantes, mesmo quando elas divergem do padrão normativo. Já em relação às variáveis em que a clara divisão social reflete o efeito histórico do contato entre línguas, o peso do estigma social é evidente. Essas correlações são altamente reveladoras do modo de operar da ideologia linguística dominante, que lança mão de um mecanismo racista para legitimar as relações de dominação econômica e a exclusão social.

Por outro lado, essas mesmas correlações oferecem a chave para a interpretação do quadro sociolinguístico, não apenas das comunidades rurais afro-brasileiras isoladas, mas de todas as variedades populares do português brasileiro. Nesse sentido, é interessante se distinguir os aspectos da morfossintaxe do português brasileiro que são marcados socialmente daqueles que não o são. No primeiro caso, estariam: a concordância nominal e a concordância verbal, e a flexão de caso dos pronomes pessoais; no segundo caso, estariam: a realização do sujeito pronominal, as relativas cortadoras, as estruturas de tópico e foco e de negação. O

uso do indicativo pelo imperativo, do pronome *ele* como acusativo e da expressão *a gente* como pronome da 1ª pessoa do plural seriam levemente marcadas do ponto de vista da avaliação social, enquanto a alternância dativa, não obstante seja estrutura mais radicalmente divergente, não sofreria, em princípio, marcação explícita, em função do desconhecimento que dela têm os falantes urbanos escolarizados. Entretanto, é possível que os falantes das comunidades afro-brasileiras possam identificar na construção do objeto duplo uma marca potencialmente estigmatizante em função do seu caráter divergente, de modo que o peso dos fatores sociais possa se fazer sentir.

A distinção em função do valor social de cada variável linguística assume um caráter decisivo para a hipótese que orientou as análises sociolinguísticas contidas neste livro. Estaria em curso um nivelamento linguístico do interior do país a partir da influência dos modelos de prestígio das grandes cidades brasileiras. Tal processo de difusão das variantes do modelo urbano culto brasileiro seria implementado principalmente pela ação dos meios de comunicação de massa e do deslocamento populacional, com a contribuição da ampliação do sistema público de educação formal. Tal cenário sociolinguístico prevê que estejam em curso, nas comunidades de fala analisadas, processos de mudança linguística de cima para baixo nos quais a variante valorizada socialmente penetra na comunidade movida por influências externas. A atualização desse processo nas análises em **tempo aparente** implica:

(i) uma curva geracional ascendente com os falantes mais jovens empregando com frequência crescente a variante inovadora (i.é, a variante de prestígio que está sendo implementada na comunidade pelas influências externas);

(ii) uma maior frequência de emprego dessa variante inovadora na fala dos indivíduos com um nível maior de contatos externos (a observação empírica realizada isolou os indivíduos que já tinham vivido pelo menos seis meses fora da comunidade);

(iii) a liderança do processo de mudança por parte dos homens, porque, no universo cultural observado, eles têm mais contato com o mundo exterior do que as mulheres, no geral presas ao universo doméstico e do trabalho na roça — um cenário semelhante foi observado por Bortoni-Ricardo (1985) e Rodrigues (1992) para a periferia das grandes cidades brasileiras;

(iv) uma frequência superior de uso da variante de prestígio por parte dos falantes com algum contato com a escolarização, mesmo que esse contato seja mínimo.

Como já foi dito, os cenários sociolinguísticos que mais se aproximaram desse arquétipo foram aqueles descritos para as variáveis linguísticas marcadas socialmente. Na concordância verbal junto à 1ª pessoa do singular (*e.g.*, **eu trabalho** *versus* **eu trabalha**), observou-se um padrão ascendente, com os jovens empregando

mais a variante padrão que os mais velhos; o diagnóstico da mudança em progresso foi confirmado com os resultados das variáveis sexo e estada fora da comunidade: os homens e aqueles que viveram fora da comunidade lideram o processo de mudança em favor da variante padrão. Entretanto, a escolaridade apresentou um resultado inconclusivo.[11] A concordância verbal com a 1ª pessoa do plural (*e.g.*, **nós trabalhamos** *versus* **nós trabalha**) apresentou praticamente o mesmo padrão de mudança em favor da variante padrão, conquanto o resultado da variável faixa etária não tenha sido ratificado pelo Programa das Regras Variáveis; o que pode ser explicado pela concorrência do processo de substituição do pronome sujeito *nós* pela forma *a gente*. O diagnóstico de mudança em favor da variante padrão no que concerne à concordância verbal junto à 3ª pessoa do plural (*e.g.*, **eles trabalham** *versus* **eles trabalha**) fundamentou-se nos resultados das variáveis faixa etária e sexo; os resultados das variáveis estada fora da comunidade e escolaridade não tiveram valor estatístico. No que concerne à concordância de número e gênero no interior do SN (*e.g.*, **meus irmão** e **no roça**, respectivamente), encontrou-se um padrão praticamente idêntico ao da concordância verbal com a 1ª pessoa do singular: um processo de mudança em curso de implementação da regra de concordância liderado pelos homens mais jovens que viveram algum tempo fora da comunidade.[12] Já os resultados das variáveis sociais da concordância nominal de gênero nas estruturas da voz passiva e de predicativo do sujeito (e.g., **essas planta já foi colhido** e **as coisa tá caro**, respectivamente) foram menos conclusivos, pois apenas o padrão geracional ascendente obteve valor estatístico; o resultado da variável sexo foi inconclusivo e a liderança dos homens e dos semianalfabetos foi indicada apenas pelos valores percentuais, não sendo selecionadas como estatisticamente relevantes pelo VARBRUL. Por fim, a alternância dativa, que pode ser incluída com ressalvas entre essas variantes socialmente marcadas também apresentou resultados indicadores de mudança em curso (descrioulizante, neste caso), com destaque para o padrão ascendente de desaparecimento da construção de objeto duplo (e.g. **dei os meninos o remédio** por **dei o remédio aos meninos**) impulsionado pela escolarização; entretanto, as variáveis sexo e estada fora da comunidade não apresentaram resultados conclusivos. A única variável linguística marcada socialmente, para a qual não se encontrou um perfil sociolinguístico

[11] Os resultados dos pesos relativos para essa variável devem ser vistos como o resultado do enviesamento decorrente da superposição da variável faixa etária sobre essa variável. Os falantes mais jovens são, em sua maioria, escolarizados, enquanto a maioria dos mais velhos é analfabeta. Quando o VARBRUL cruza as duas variáveis, inverte os pesos dos fatores ligados à escolaridade, pois essa variável é selecionada depois da faixa etária, à qual o programa atribui um maior peso.

[12] Poder-se-ia acrescentar a característica de terem algum contato com a escolaridade, conquanto os resultados dessa variável tenham sido comprometidos em função do desequilíbrio nas células na composição das amostras de fala (cf. nota anterior).

definido foi a flexão de caso dos pronomes pessoais, em função de uma série de dificuldades intrínsecas, como a baixa frequência de ocorrências dessa variável em função da natureza das entrevistas realizadas, relacionada com o fato de a variação que ainda se observa, com formas desviantes do tipo **viu eu**, **gosta de tu** e **terra de nós**, ter atualmente um caráter quase vestigial.[13]

Considerando-se todas as dificuldades em se estabelecer correlações entre os processos linguísticos e os sociais que ainda persistem no estágio atual do desenvolvimento da pesquisa sociolinguística em geral e no Brasil em particular, esses resultados podem ser tidos como muito elucidativos,[14] e permitem reunir um conjunto representativo de mudanças em direção ao padrão urbano que ratificam a hipótese de um nivelamento linguístico em curso no interior do país, no bojo do qual estariam se esgarçando as marcas mais características do contato entre línguas que caracteriza a formação histórica do português popular brasileiro. Tomando como parâmetro as variáveis sociais aqui analisadas, constata-se:

(i) o padrão ascendente em favor da variante do padrão urbano para praticamente todas as variáveis linguísticas marcadas socialmente;

(ii) que os homens lideram esse processo de mudança em cinco das sete variáveis analisadas (em duas o resultado dessa variável é inconclusivo);

(iii) que aqueles que viveram pelo menos seis meses fora das comunidades usam a variante padrão com mais frequência igualmente em cinco das variáveis analisadas, e, em uma, sem confirmação estatística;

(iv) que a variável escolaridade foi a que apresentou resultados mais difusos, o que estaria relacionado com a precariedade do sistema de educação pública que atende às comunidades estudadas.[15]

O Quadro 1 apresenta, de forma esquemática, o conjunto dos condicionamentos sociais dessas variáveis linguísticas socialmente marcadas.

Deve-se destacar, em primeiro lugar, que não houve, para esse conjunto de variáveis linguísticas socialmente marcadas, qualquer resultado que contrariasse o diagnóstico de uma mudança em curso em favor da aquisição das marcas morfológicas presentes no padrão urbano culto. Por outro lado, tais resultados refutam a hipótese de uma deriva secular defendida por Naro e Scherre (1993, 2007), pois essa hipótese implica um processo gradual de perda das marcas morfológicas através de um sem número de gerações. É de fato surpreendente que

---

[13] Pesquisas estão em curso na tentativa de superar tais dificuldades e de estabelecer um perfil sociolinguístico para essa variável.

[14] Pode-se mesmo afirmar que não há ainda um conjunto de resultados tão consistentes para qualquer outra variedade social do português brasileiro.

[15] Devendo-se ter em conta também os problemas metodológicos já referidos.

VARIÁVEIS MUITO MARCADAS	PERFIL				
	Idade	**Sexo**	**Estada fora**	**Escolaridade**	**Comunidade**
Concordância verbal com a 1ª pessoa do singular	**padrão ascendente**	**homens lideram**	**confirmado**	resultado inconclusivo	variação só em Helvécia
Concordância verbal com a 1ª pessoa do plural	padrão ascendente (sem valor estatístico)	**homens lideram**	**confirmado**	sem valor estatístico	sem valor estatístico
Concordância verbal com a 3ª pessoa do plural	**padrão ascendente**	**homens lideram**	sem valor estatístico	sem valor estatístico	Cinzento mais divergente (que Helvécia e Rio de Contas)
Concordância de gênero em predicativos e passivas	**padrão ascendente**	resultado inconclusivo	confirmado apenas nos percentuais	confirmado apenas nos percentuais	sem resultado
Concordância de número no Sintagma Nominal[1]	**padrão ascendente**	**homens lideram**	**confirmado**	resultado inconclusivo	análise restrita a Helvécia
Concordância de gênero no Sintagma Nominal	**padrão ascendente**	**homens lideram**	**confirmado**	confirmado apenas nos percentuais	variação só em Helvécia
Alternância dativa	**padrão ascendente**	sem valor estatístico	sem valor estatístico	**confirmado**	RC e HV mais divergentes (seguidas de Cinzento e Sapé)
Flexão de caso	não observada	não observada	não observada	não observada	não observada

Quadro 1 - Condicionamentos sociais das variáveis linguísticas marcadas por um julgamento subjetivo explícito

[1] Com base nos resultados de Andrade (2003).

dois dos maiores sociolinguistas do país defendam uma hipótese para a qual não se pode encontrar qualquer evidência empírica nas análises sociolinguísticas das diversas variedades do português brasileiro. A tendência aqui observada de mudança aquisicional de marcas morfológicas encontra paralelo em diversas análises sociolinguísticas do português popular brasileiro (NINA, 1980; BORTONI-RICARDO, 1985; RODRIGUES, 1992; VIEIRA, 1997, entre outros). E, para o português urbano, os resultados das variáveis sociais da concordância verbal e nominal reunidos por esses mesmos autores aponta para um quadro de variação estável (SCHERRE; NARO, 1997). Portanto, em nenhum dos casos, identifica-se um processo de mudança em curso de perda gradual de marcas flexionais que pudesse ser a atualização de uma deriva secular da língua nessa direção.

Mas, se o cenário de uma mudança aquisional é nítido nos processos de variação decorrentes da erosão gramatical provocada pelo contato entre línguas, não se encontrou nas comunidades de fala estudadas uma distribuição clara que indicasse um processo de mudança em curso na análise sociolinguística das variáveis para as quais o efeito do contato entre línguas não é muito nítido. Uma das razões para isso seria o fato de essas variáveis não serem marcadas com um julgamento social explícito. Esse seria o caso dos processos de variação na realização do sujeito pronominal, nas construções relativas e nas estruturas de negação, e dos processos de topicalização e clivagem. Como se pode ver no Quadro 2, quase não há distribuição social visível nos processos de variação e mudança.

Na análise de duas dessas variáveis, os fatores sociais não foram sequer considerados. Nas outras três, a realização do sujeito pronominal não apresentou um padrão definido na variável faixa etária, vislumbrou-se um padrão descendente com relação à topicalização, que seria mais frequente na fala dos mais velhos, e um quadro de variação estável no processo de variação envolvendo as estruturas de negação. No caso da realização do sujeito pronominal, os dois únicos resultados estatisticamente válidos das variáveis sociais (sexo e estada fora da comunidade) foram contraditórios entre si, pois o sujeito nulo predomina entre os homens e entre aqueles que nunca viveram fora da comunidade. No caso da topicalização, a análise não apresenta outras evidências que comprovem uma tendência à diminuição do uso de estruturas topicalizadas. Por fim, o quadro de variação estável nas estruturas de negação se confirma apenas com o resultado da variável social estada fora da comunidade, em que a variante padrão da negação pré-verbal predomina na fala daqueles que viveram fora da comunidade.

Já no caso das variáveis que, mesmo não sendo diretamente relacionadas com o contato entre línguas, sofrem um julgamento social em algum nível, os efeitos dos condiconamentos sociais tornam-se um pouco mais nítidos, como se pode ver no Quadro 3.

VARIÁVEIS NÃO MARCADAS	PERFIL				
	Idade	Sexo	Estada fora	Escolaridade	Comunidade
Sujeito nulo	sem valor estatístico	**homens lideram**	**padrão invertido**	sem valor estatístico	sem resultado
Orações relativas	não observada	não observada	não observada	não observada	não observada
Clivagem	não observada	não observada	não observada	não observada	não observada
Topicalização	**padrão descendente**	não observada	não observada	**não interfere**	não observada
Negação	**variação estável**	sem valor estatístico	**confirmado**	sem valor estatístico	não observada

Quadro 2 - Condicionamentos sociais das variáveis linguísticas não marcadas por um julgamento subjetivo explícito

VARIÁVEIS POUCO MARCADAS	PERFIL				
	Idade	Sexo	Estada fora	Escolaridade	Comunidade
Uso do artigo no SN	**padrão ascendente**	não observada	não observada	não observada	Análise restrita a Helvécia
Uso do modo Subjuntivo	**padrão ascendente**	sem valor estatístico	sem valor estatístico	**confirmado**	Helvécia mais divergente (seguida de CZ, RC e SP)
Objeto Direto Anafórico	inconclusivo	inconclusivo	inconclusivo	inconclusivo	inconclusivo
Nós *versus* a gente	**ascendente para *a gente***	sem valor estatístico	sem valor estatístico	**“invertido”**	CZ mais conservadora (seguida de RC, HV e SP)
Nosso *versus* da gente	**ascendente para *da gente***	sem valor estatístico	**mudança em favor de *da gente***	sem valor estatístico	CZ mais conservadora (seguida de RC, HV e SP)

Quadro 3 - Condicionamentos sociais das variáveis linguísticas com algum julgamento subjetivo

O emprego do artigo definido apresentou um padrão ascendente na variação geracional, com os jovens usando mais o artigo definido que os mais velhos. É bem provável que se encontrassem outros indícios desse processo de mudança aquisicional, mas, infelizmente, a análise contida neste livro não explorou as demais variáveis sociais. O emprego do modo subjuntivo também apresentou um padrão geracional de mudança no sentido da substituição das formas do indicativo pelas formas do subjuntivo nos contextos hipotéticos, contrafactuais e de incerteza. O diagnóstico de mudança de cima para baixo foi confirmado pela variável escolaridade, com o uso mais frequente do modo subjuntivo entre os falantes com alguma experiência de escolarização, porém os resultados das variáveis sexo e estada fora da comunidade não lograram confirmação estatística. No caso do objeto direto anafórico de 3ª pessoa, os resultados sociais foram inconclusivos, já que a análise apenas vislumbrou um processo de incremento do pronome lexical em contextos da categoria vazia. A pressão externa, nesse caso, é quase nula, na medida em que a variante do padrão normativo, o clítico acusativo, está desaparecendo da fala até mesmo nos segmentos urbanos escolarizados. Já no processo de implementação da forma *a gente* em substituição do pronome canônico *nós* (inclusive na forma do possessivo), os condicionamentos sociais identificados parecem apontar para um processo de mudança de baixo para cima, em que a variante inovadora vai se expandindo nos diversos níveis da interação linguística sub-repticiamente, sem a força modeladora do **prestígio**, no sentido que lhe foi conferido por Labov (1972a, 1982).

Por outro lado, ainda no plano da caracterização sócio-histórica do português afro-brasileiro, e cotejando os resultados das quatro comunidades que constituem o universo de observação desta pesquisa, constatou-se que, apesar de suas características comuns, o português afro-brasileiro, como qualquer outra variedade linguística, não constitui um universo homogêneo. Nesse sentido, confirmou-se a expectativa inicial, com a comunidade de Helvécia, para a qual se tem algum registro de um passado crioulizante, sendo aquela que apresentou os padrões coletivos de comportamento linguístico mais divergentes em relação ao padrão urbano culto. Houve processos de variação que só foram observados nessa comunidade: os que afetam a concordância de gênero no interior do SN e a concordância verbal com a 1ª pessoa do singular. Considerando-se a estrutura identificada como a mais crioulizante nesta pesquisa, as construções com objeto duplo foram mais frequentes em Helvécia e Rio de Contas. Helvécia também exibiu a menor frequência de emprego das formas do subjuntivo. A comunidade de Cinzento foi a que apresentou a menor frequência de concordância verbal junto à 3ª pessoa do plural; foi também a mais refratária à substituição de *nós* pela expressão *a gente*, seguida por Rio de Contas. E, no que concerne a essa variável, Helvécia aproximou-se mais de Sapé, a comunidade que mais incorporou os padrões urbanos, em

praticamente todas as variáveis linguísticas que foram analisadas. Isso pode ser explicado pela maior proximidade de Sapé em relação à capital do Estado, a cidade do Salvador, grande centro irradiador de modelos linguísticos e culturais da região. Portanto, pode-se colocar Helvécia e Sapé, nos dois extremos de uma escala em que se dispõem as quatro comunidades estudadas, com Cinzento e Rio de Contas ocupando posições intermediárias.

O conflito que se observou entre os resultados obtidos só vem a confirmar a impossibilidade de estabelecer correlações lineares e mecânicas, em função da complexidade da rede inextrincável de mediações que se interpõem nas relações entre a estruturação linguística e os processos históricos e sociais. A complexidade que se identificou no universo restrito das comunidades rurais afro-brasileiras dá uma pequena ideia da dimensão do desafio que será estabelecer relações empiricamente motivadas que atestem as relações históricas do contato entre línguas na formação da realidade linguística brasileira. E isso só vem a ratificar o significado dos resultados empíricos reunidos neste livro como forma de lançar luzes sobre a história sociolinguística do Brasil.

Palavras finais

No plano mais geral da história sociolinguística do Brasil, a maior contribuição deste livro foi a de ter sistematizado um conjunto de análises que revelaram como o contato entre línguas afetou particularmente uma variedade da língua portuguesa transplantada para o continente americano; variedade esta aqui individualizada e denominada **português afro-brasileiro**. Nesse sentido, este volume resgata uma rica relação que ligava o Brasil ao continente africano até meados do século XIX, relação esta desbaratada pela ação do imperialismo britânico (COSTA E SILVA, 2003). Por outro lado, retrata como as desigualdades da sociedade brasileira se refletem no plano da língua, configurando a polarização sociolinguística do país, que tem profundas raízes históricas, nas quais a integração na sociedade brasileira de milhões de africanos escravizados ao longo de mais de três séculos desempenhou um papel decisivo. Nesse amplo cenário sociolinguístico do Brasil de hoje, os padrões coletivos de comportamento linguístico das comunidades rurais afro-brasileiras isoladas resultam de dois processos sócio-históricos contrários, que acabam por definir as feições de toda a norma popular brasileira. Por um lado, estão as forças invisíveis de um passado obscuro do qual os efeitos do contato ainda se fazem sentir como ondas que se propagam sob a superfície do mar. Por outro lado, estão as mudanças que se expandem com a difusão dos padrões urbanos para todas as regiões do país, como as correntes de vento que definem a direção das ondas na superfície do mar.

Portanto, pode-se pensar em dois tipos de mudança que estariam em curso dentro das comunidades de fala analisadas. De um lado, estariam as mudanças que decorreriam de pressões estruturais invisíveis e/ou de variantes inovadoras que estariam alterando o padrão linguístico de forma subliminar, em um processo que ocorre abaixo do nível de consciência (LABOV, 1972a, 1982). Nesse caso se encontram as mudanças que não estariam diretamente ligadas ao contato entre línguas (como a substituição de *nós* pela expressão *a gente*), bem como mudanças com origem no contato entre línguas que se teriam difundido para todas as variedades do português brasileiro, não sendo avaliadas socialmente de forma negativa. Esse poderia ser o caso da realização do sujeito pronominal, no qual a perda do sujeito nulo poderia estar se processando, de forma mais ou menos uniforme, em todas as variedades do português brasileiro. Pesquisas futuras sobre as orações relativas poderão deslindar um cenário análogo, em que a perda de orações relativas com antecedente introduzidas por relativizadores pronominais tenha sido uma mudança desencadeada pelo contato que se espraiou para todas as variedades do português brasileiro.

Por outro lado, as mudanças estariam embutidas no processo de nivelamento linguístico em curso no país a partir dos modelos urbanos cultos. Nesse caso, processos de mudança de cima para baixo estariam em curso nas comunidades aqui analisadas, com a implementação das variantes de prestígio em detrimento daquelas que teriam resultado do processo de transmissão linguística irregular desencadeado pelo contato entre línguas no passado. Nesse contexto, não se encontra na comunidade qualquer sentimento de resistência cultural, que Labov (1994) chamou de **orientação para grupo**, e que começa a ganhar corpo na periferia das grandes cidades brasileiras. A orientação é para o **prestígio**, mesmo que de forma pouco consciente, pois o que vem de fora é geralmente avaliado de forma positiva como parte integrante da chegada do progresso, que trará a melhoria nas condições materiais tão adversas da vida no passado, sujeita às intempéries do clima, sem qualquer assistência médica, sem comunicação e sem acesso a meios de locomoção dignos. As velhas formas provenientes do contato linguístico são assim como fantasmagorias de um passado remoto e sombrio, onde se encaverna a chaga da escravidão; passado esse que se quer esquecer.

O estudo das outras variedades populares do interior do país e das grandes cidades, que constituem as etapas futuras desta pesquisa, poderá ampliar os horizontes desse cenário, identificando o vetor em direção ao padrão urbano culto que caracteriza a norma popular brasileira atualmente. Por outro lado, o cotejo entre outras variedades populares e as comunidades aqui analisadas poderá trazer novas evidências empíricas que ratifiquem, alterem ou ampliem os achados desta pesquisa, que focaliza os efeitos do contato entre línguas na história sociolinguística do país. Porém, os futuros avanços não anularão o valor dos resultados alcançados

até aqui no sentido de resgatar a dimensão pluriétnica da sociedade brasileira no plano da sua realidade linguística. Para além dos avanços teoréticos no campo da pesquisa sobre o contato linguístico, sobressai a mudança de foco da pesquisa histórica: ao invés da análise da evolução interna da língua da elite, a visão ampla da complexidade histórica da formação dos padrões linguísticos da grande maioria da população brasileira, em toda a sua diversidade étnica e cultural. E, sobretudo, o resgate da fala daqueles que nunca tiveram vez nem voz e se refugiaram nas entranhas mais recônditas do país para alcançar a liberdade. E dessa fala esquecida recolher os elementos mais preciosos para a compreensão do que está no cerne de todas as línguas portuguesas do Brasil.

Referências

ABNEY, Steven. *The English noun phrase in its sentential aspect*. 1987. Tese (Doutorado em Linguística) - Massachusetts Institute of Technology, Cambridge.

ADAMS, Marianne Patalino. *Old French, null subjects, and verb second phenomena*. Tese (Doutorado em Linguística) - Universidade da Califórnia, Los Angeles.

AIKHENVALD, Alexandra. *Language contact in Amazonia.* Oxford: Oxford University Press, 2002.

ALEXANDRE, Nélia Maria Pedro. *A estratégia resumptiva em relativas restritivas do português europeu*. 2000. Dissertação (Mestrado) - Universidade de Lisboa, Lisboa.

ALEXANDRE, Nélia Maria Pedro; HAGEMEIJER, Tjerk. *Pronomes resumptivos e abandono de preposição nos crioulos atlânticos de base portuguesa*. Universidade Autónoma de Lisboa; Universidade de Lisboa, 2001. Ms.

ALEXANDRE, Nélia; HAGEMEIJER, Tjerk. Bare nouns and the nominal domain in Santome. In: BATISTA, Marlyse; GUERÓN, Jaqueline (Ed.). *Noun Phrase in Creole languages*: *a multi-faceted approach*. Amsterdam: John Benjamins, 2007. p.37-59.

ALKMIM, Mônica Guieiro Ramalho de. Negativa pré e pós-verbal: implementação e transição. In: COHEN, Maria Antonieta A. M.; RAMOS, Jânia M. (Org.). *Dialeto mineiro e outras falas*. Belo Horizonte: Faculdade de Letras UFMG, 2002. v.1. p.169-182.

ALMADA, Maria Dulce de Oliveira. *Cabo Verde*: contribuição para o estudo do dialeto falado no seu arquipélago. Lisboa: Junta de Investigações do Ultramar, 1961.

ALMEIDA, Alessandra. *A concordância verbal na comunidade de São Miguel dos Pretos, Restinga Seca, RS*. 2005. Dissertação (Mestrado em Estudos da Linguagem) - Universidade Federal do Rio Grande do Sul, Porto Alegre.

AMARAL, Amadeu. *O dialecto caipira*. São Paulo: "O Livro", 1920.

ANDERSEN, Roger. The dimensions of pastness. In: SALABERRY, Rafael; SHIRAI, Yasuhiro (Ed.). *The L2 acquisition of tense-aspect morphology*. Amsterdam: John Benjamins, 2002. p.79-105.

ANDRADE, P. Ribeiro de. *Um fragmento da constituição sócio-histórica do português do Brasil*: variação na concordância nominal de número em um dialeto afro-brasileiro. 2003. Dissertação (Mestrado em Linguística) - Universidade Federal da Bahia, Salvador.

ANJOS, Rafael S. A. *Territórios das comunidades remanescentes de antigos quilombos no Brasil*: primeira configuração espacial. 2.ed. Brasília: Mapas Editora & Consultoria, 2000.

ARAÚJO, Edivalda A. *As construções de tópico do português dos séculos XVIII e XIX*: uma análise sintático-discursiva. 2006. Tese (Doutorado em Linguística) - Universidade Federal da Bahia, Salvador.

ARIM, Eva; RAMILO, Maria Celeste; FREITAS, Tiago. Estratégias de relativização nos meios de comunicação social portugueses. In: ENCONTRO DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA, 19. *Actas*... Lisboa: APL; ILTEC, 2004. Disponível em: <http://www.iltec.pt/pdf/wpapers/2005-redip-relativas.pdf>. Acesso em: 23 nov. 2007.

AVERBUG, Mayra C. G. *Objeto direto anafórico e sujeito pronominal na escrita de estudantes*. 2000. Dissertação (Mestrado em Letras) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

AZEVEDO, Thales. *Democracia racial*. Petrópolis: Vozes, 1975.

BAGNO, Marcos. *Português ou brasileiro? Um convite à pesquisa*. São Paulo: Parábola, 2001.

BAHIA. Governo do Estado. *A inserção da Bahia na evolução nacional - 1ª etapa, 1850-1889*: atividades produtivas. v.2. Salvador: Secretaria do Planejamento, Ciência e Tecnologia, 1978.

BAIOCCHI, Mari de Nasaré. *Kalunga*: *estórias e textos*. Goiânia: Secretaria de Estado da Educação de Goiás, 1991.

BAKER, Marc C. Thematic roles and syntactic structure. In: HAEGMAN, Liliane. (Ed.). *Elements of grammar*: handbook of generative syntax. Dordrecht: Kluwer, 1997. p.73-137.

BAKER, Philip. On the origins of the first Mauritians and of the creole languages of their descendants. In: BAKER, Philip; CORNE, Chris. *Isle de France Creole*. Ann Arbor: Karoma, 1982. p.135-163.

BAO, Z. The aspectual system of Singapore English and the systemic substratist explanation. *Journal of Linguistics*, Cambridge, v.41, Issue 2, p.237-267, 2005.

BAPTISTA, Marlyse. *The syntax of Cape Verdean Creole*: the Sotavento varieties. Amsterdam: John Benjamins, 2002.

BAPTISTA, Marlyse. On the syntax and semantics of DP in Cape Verdean Creole. In: BAPTISTA, Marlyse; GUÉRON, Jacqueline (Ed.). *Noun phrases in creole languages*: a multi-faceted approach. Amsterdam; Philadelphia: John Benjamins, 2007. p.61-105.

BAPTISTA, Marlyse; GUÉRON, Jacqueline (Ed.). *Noun phrases in creole languages*: a multi-faceted approach. Amsterdam; Philadelphia: John Benjamins, 2007.

BARBOSA, Pilar; KATO, Mary A.; DUARTE, M. Eugênia L. A distribuição do sujeito nulo no português europeu e no português brasileiro. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO PORTUGUESA DE LINGUÍSTICA, 16., 2001, Lisboa. *Actas*... Lisboa: APL, 2001. p.539-550.

BAXTER, Alan. A contribuição das comunidades afro-brasileiras isoladas para o debate sobre a crioulização prévia: um exemplo do Estado da Bahia. In: D'ANDRADE, Ernesto; KIHM, Alain (Org.). *Actas do Colóquio sobre Crioulos de Base Lexical Portuguesa*. Lisboa: Colibri, 1992. p.7-36.

BAXTER, Alan. Portuguese and Creole Portuguese in the Pacific and Western Pacific rim. In: WURM, P.; MÜHLHÄUESLER, D. T. Tryon (Ed.). *Atlas of languages of intercultural communication in the Pacificf, Asia, and the Americas*. Berlin: Mouton, 1996. p.301-338.

BAXTER, Alan. O português vernáculo do Brasil - Morfossintaxe. In: PERL, Matthias; SCHWEGLER, Armin (Ed.). *América negra*: panorámica actual de los estudios

lingüísticos sobre variedades hispanas, portuguesas y criollas. Frankfurt am Main: Vervuert; Madrid: Iberoamericana, 1998. p.97-134.

BAXTER, Alan. The context of language acquisition among slaves of the Colônia Leopoldina. Trabalho apresentado ao Research Seminar of the Institute of Latin American Studies, La Trobe University, 1999.

BAXTER, Alan. Semicreolization? The restructured Portuguese of the Tongas of São Tomé, a consequence of L1 acquisition in a special contact situation. *Journal of Portuguese Linguistics*, Lisboa, v.1, n.1, p.7-39, 2002.

BAXTER, Alan. A re-estruturação do português numa situação de contato africana. Trabalho apresentado ao 3º Congresso Internacional da Associação Brasileira de Linguística, Rio de Janeiro, 2003.

BAXTER, Alan. The development of variable NP plural agreement in a restructured African variety of Portuguese. In: ESCURE, G.; SCHWEGLER, A. (Ed.). *Creoles, contact and language change*: linguistics and social implications. Amsterdam: John Benjamins, 2004. p.97-126.

BAXTER, Alan; LOPES, Norma. Variação no uso de artigos: uma comparação. Trabalho apresentado à 20ª Jornada Nacional de Estudos Linguísticos, João Pessoa, 2004.

BAXTER, Alan; LOPES, Norma. Bare definite reference NPs in an Afro-Brazilian Portuguese dialect. *Revista Internacional de Lingüística Iberoamericana*, Madrid, v.4, n.1, p.55-70, 2006.

BAXTER, Alan; LUCCHESI, Dante. Processos de descrioulização no sistema verbal de um dialeto rural brasileiro. *Papiá - Revista de Crioulos de Base Ibérica*, Brasília, Universidade de Brasília, v.2, n.2, p.59-71, 1993.

BAXTER, Alan; LUCCHESI, Dante. A relevância dos processos de pidginização e crioulização na formação da língua portuguesa no Brasil. *Estudos Linguísticos e Literários*, Salvador, n.19, p.65-83, 1997.

BAXTER, Alan; LUCCHESI, Dante. Un paso más hacia la definición del pasado criollo del dialecto afro-brasileño de Helvécia (Bahia). In: ZIMMERMANN, Klaus (Ed.). *Lenguas criollas de base lexical española y portuguesa*. Madri: Iberoamericana, 1999. p.119-141.

BAXTER, Alan; LUCCHESI, Dante; GUIMARÃES, Maximiliano. Gender agreement as a decreolizing feature of the Afro-Brazilian rural dialect of Helvécia. *Journal of Pidgin and Creole Languages*, Filadélfia; Amsterdam, v.12, n.1, p.1-57, 1997.

BELAZI, Heidi. *Multilingualism in Tunisia and code-switching among educated Tunisian bilinguals*. 1991. Tese (Doutorado em Linguística) - Cornell University, Ithaca.

BELAZI, Heidi; RUBIN, Edward; TORIBIO, Almeida. Code-switching and X-bar theory: the functional head constraint. *Linguistic Inquiry*, Massachusetts, v.25, n 2, p.221-237, 1994.

BELLETTI, Adriana. Aspects of the low TP area. In: RIZZI, L. (Ed). *The structure of CP and IP. The cartography of syntactic structures*. v.2. New York: Oxford University Press, 2004. p.16-51.

BELLETTI, Adriana. Answering with a cleft: the role of the null subject parameter and the VP perihery. In: BRUGÈ, L. et al. (Ed.). *Proceedings of the Thirtieth "Incontro di Grammatica Generativa"*. Veneza: Cafoscarina, 2005. p.63-82.

BENVENISTE, Emile. A natureza dos pronomes. In: DASCAL, Marcel (Org.). *Fundamentos metodológicos da linguística*. v.4: Pragmática - Problemas, críticas, perspectivas da Linguística. Campinas: Ed. do Autor, 1982.

BIBERAUER, Theresa; RICHARDS, Marc. True optionality: when the grammar doesn't mind. In: BOECKX, C. (Ed.). *Minimalist theorizing*. Amsterdam: John Benjamins, 2006. p.35-67.

BIBERAUER, Theresa; ROBERTS, Ian. Changing EPP-parameters in the history of English: accounting for variation and change. *English Language and Linguistics*, Cambridge, v.9, n.1, p.1-42, 2005.

BICKERTON, Derek. *Roots of language*. Ann Arbor: Karoma, 1981.

BICKERTON, Derek. The Language Bioprogram Hypothesis. *Behavioural and Brain Sciences*, Cambridge, n.7, 1984, p.173-203.

BICKERTON, Derek. Creole languages and the Bioprogram. In: NEWMEYER, Frederick (Ed.). *Linguistics*: the Cambridge survey. v.2. Cambridge: Cambridge University Press, 1988. p.268-284.

BICKERTON, Derek. How to acquire language without positive evidence: what acquisitionists can learn from Creoles? In: DEGRAFF, Michel (Ed.). *Language creation and language change*: creolization, diachrony, and development. Cambridge: The MIT Press, 1999. p.49-74.

BOLE-RICHARD, Rémy. *Systématique phonologique et gramaticale d'un parler éwé*: le genmina du Sud-Togo et Sud-Bénin. Paris: L'Harmattan, 1983.

BORER, Hagit. *Parametric syntax*. Dordrecht: Foris, 1984.

BORER, Hagit. Anaphoric AGR. In: JAEGGLI, O.; SAFIR, K. J. (Ed.). *The null subject parameter*. Dordrecht: Kluwer, 1989. p.69-110.

BORTONI-RICARDO, Stella Maris. *The urbanization of rural dialect speakers*: a sociolinguistic study in Brazil. Cambridge: Cambridge University Press, 1985.

BRAGA, Maria Luísa. As sentenças clivadas no português falado do Rio de Janeiro. *Organon*, Porto Alegre, v.5, n.18, p.109-125, 1991.

BRANDÃO, Sílvia Figueredo; MOTA, Maria Antónia. *Análise contrastiva de variedades do português*. Rio de Janeiro: In-Fólio, 2003.

BRITO, Ana Maria Barros de. *A sintaxe das orações relativas em português*. Porto: Instituto Nacional de Investigação Científica, 1991.

BRITO, Ana Maria; DUARTE, Inês. Orações relativas e construções aparentadas. In: MATEUS, M. H. Mira et al. *Gramática da língua portuguesa*. 5.ed. revista e aumentada. Lisboa: Caminho, 2003. p.653-694.

BRITO, Ana Maria; DUARTE, Inês; MATOS, Gabriela. Frases com tópicos marcados. In: MIRA MATEUS et al. *Gramática da língua portuguesa*. 5.ed. rev. e aum. Lisboa: Caminho, 2003. p.489-502.

BRUHN DE GARAVITO, Joyce. Learner's competence may be more accurate than we think: Spanish L2 and agreement morphology. In: LICERAS, Juana M.; ZOBL, Helmut; GOODLUCK, Helen (Ed.). *Proceedings of the 6th Generative Approaches to Second Language Acquisition.* Somerville: Cascadilla Proceedings Project, 2003. p.17-23.

BRUYN, Adrienne; MUYSKEN, Pieter; VERRIPS, Maaike. Double-object constructions in the creole languages: development and acquisition. In: DEGRAFF, Michel (Ed.). *Language creation and language change*: creolization, diachrony, and development. Cambridge: The MIT Press, 1999. p.329-374.

BUESCO, Mircea. *Exercícios de história econômica do Brasil*. Rio de Janeiro: Padrão, 1968.

BURGOS, Luís Eduardo S. de. *Estratégias de uso das relativas em uma comunidade de fala afro-brasileira*. 2003. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal da Bahia, Salvador.

CACCIAGUERRA, Vanessa. *O emprego do 'onde' no português paulistano culto atua*l. 2006. Dissertação (Mestrado em Filologia e Língua Portuguesa) - Universidade de São Paulo, São Paulo.

CAIN, J; WEBER-OLSEN, M; SMITH, R. Acquisition strategies in a first and second language: are they the same? *Journal of Child Language*, Cambridge, v.14, n.2, p.333-352, Jun. 1987.

CALÓGERAS, J. Pandiá. *Formação histórica do Brasil*. 5.ed. São Paulo: Nacional, 1957 [1927].

CÂMARA JR., Joaquim Mattoso. Ele como um acusativo no português do Brasil. In: ____. *Dispersos*. Rio de Janeiro: Fundação Getúlio Vargas, 1972. p.47-54.

CÂMARA JR., Joaquim Mattoso. *História e estrutura da língua portuguesa*. Rio de Janeiro: Padrão, 1979.

CÂMARA JR., Joaquim Mattoso. *Estrutura da língua portuguesa*. Petrópolis: Vozes, 2002 [1970].

CAMARGOS, Marcelo. *A negativa*: uma análise qualitativa, 2000. Disponível em <http://www.ufop.br/ichs/conifes/anais/LCA/clca03.htm>. Acesso em: 30 jul. 2005.

CARDOSO, Eduardo Augusto. *O crioulo da Ilha de São Nicolau de Cabo Verde*. Lisboa: Instituto de Cultura e Língua Portuguesa; Praia: Instituto Caboverdiano do Livro, 1990.

CARENO, Mary Francisca do. *A linguagem rural do Vale do Ribeira*: a voz e a vez das comunidades negras. 1991. Tese (Doutorado em Filologia e Linguística Portuguesa) - Universidade Estadual Paulista, Assis.

CARENO, Mary Francisca do. As narrativas fantásticas do negro rural: a resistência do cotidiano. In: CONGRESSO BRASILEIRO DE PESQUISADORES NEGROS, 1., 2000, *Recife*. *Anais*... Recife: Universidade Federal de Pernambuco, 2000. CD-ROM.

CARVALHO, Guilhermina M. B. Mendes de. *A inacusatividade na fala de comunidades rurais afro-brasileiras isoladas*. 2006. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal da Bahia, Salvador.

CASTRO, Ana; COSTA, João. Weak forms as X°: Prenominal possessives and preverbal adverbs in Portuguese. In: PÉREZ-LEROUX, Ana Teresa; ROBERGE, Yves (Ed.). *Romance linguistics*: theory and acquisition. Amsterdam; Philadelphia: John Benjamins, 2003. p.95-110.

CASTRO, Ana; PRATAS, Fernanda. Capeverdean DP-internal number agreement: additional arguments for a distributed morphology approach. In: COSTA, J.;

FIGUEIREDO SILVA, M. C. (Ed.). *Studies on agreement*. Amsterdam: John Benjamins, 2006. p.11-24.

CASTRO, Yeda Pessoa de. Os falares africanos na interação social dos primeiros séculos. In: MELLO, Linalda Arruda (Org.). *Sociedade, cultura & língua*. João Pessoa: Shorin, 1990. p.91-113.

CASTRO, Yeda Pessoa de. *A língua mina-jeje no Brasil*: um falar africano em Ouro Preto do século XVIII. Belo Horizonte: Secretaria de Estado da Cultura, 2002.

CAVALCANTE, Rerisson. *A negação pós-verbal no português brasileiro*: análise descritiva e teórica de dialetos rurais de afrodescendentes. 2007. Dissertação (Mestrado em Letras) - Universidade Federal da Bahia, Salvador.

CEDERGREN, Henrietta J.; SANKOFF, David. Variable rules: performance as a statistical reflection of competence. *Language*, LSA, v.50, n.2, p.333-355, 1974.

CERQUEIRA, Vicente C. A forma genitiva *dele* e a categoria de concordância (AGR) no português brasileiro. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993. p.129-160.

CERQUEIRA, Vicente C. *Primeiro o demonstrativo; depois, o artigo*: considerações sobre a aquisição de categorias funcionais. Campinas: UNICAMP, 1993. Ms.

CERQUEIRA, Vicente C. *A sintaxe do possessivo no português brasileiro*. 1996. Tese (Doutorado em Letras) - Universidade Estadual de Campinas, Campinas-SP.

CHAMBERS, J. K.; TRUDGILL, Peter. *Dialectology*. Cambridge: Cambridge University Press, 1980.

CHIAVENATO, Júlio José. *O negro no Brasil*: da senzala à Guerra do Paraguai. São Paulo: Brasiliense, 1980.

CHIERCHIA, Gennaro. Reference to kinds across languages. *Natural Language Semantics*, Dordrecht, v.6, n.4, p.339-405, dec. 1998.

CHOMSKY, Noam. *Syntactic structures*. The Hague: Mouton, 1957.

CHOMSKY, Noam. *Aspects of the theory of syntax*. Cambridge: The MIT Press, 1965.

CHOMSKY, Noam. *Cartesian linguistics*: a chapter in the history of rationalist thought. New York: Harper & Row, 1966.

CHOMSKY, Noam. Deep structure, surface structure and semantic interpretation. In: STEINBERG, D.; JAKOBOVITS, L. (Ed.). *Semantics*: an interdisciplinary reader in philosophy, linguistics and psychology. Cambridge: Cambridge University Press, 1971. p.183-216.

CHOMSKY, Noam. *Reflections on language*. New York: Pantheon, 1975.

CHOMSKY, Noam. Conditions on rules of grammar. *Linguistic Analysis*, n.2, p.303-351, 1976.

CHOMSKY, Noam. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CHOMSKY, Noam. *Knowledge of language*: its origin, nature and use. New York: Praeger, 1986.

CHOMSKY, Noam. A minimalist program for linguistic theory. In: HALE, K.; KEYSER, S. J. (Ed.) *The view from biulding 20*. Massachusetts: MIT Press, 1993.

CHOMSKY, Noam. *The minimalist program*. Cambridge: The MIT Press, 1995.

CHOMSKY, Noam. Beyond explanatory adequacy. In: *MIT Occasional Papers in Linguistics*. Cambridge, MIT Working Papers in Linguistics, n.20. 2001a.

CHOMSKY, Noam. Derivation by phase. In: KENSTOWICZ, M. (Ed.). *Ken Hale*: a life in language. Cambridge: The MIT Press, 2001b. p.1-52.

CINQUE, Guglielmo. On the evidence for partial N-movement in the Romance DP. In: CINQUE, G. et al. (Ed.). *Paths towards Universal Grammar*. Washington, DC: Georgetown University Press, 1994. p.85-110.

CINTRA, Luís F. Lindley. *Sobre "formas de tratamento" na língua portuguesa*. Lisboa: Livros Horizonte, 1972.

CLAHSEN, Harald. Constraints on parameter setting: a grammatical analysis of some acquisition stages in German child language. *Language Acquisition*, Amsterdam, v.1, n.4, p.361-391, 1991.

CLEMENTS, J. Clancy. The Tense-Aspect system in pidgins and naturalistically learned L2. *Studies in Second Language Acquisition*, n.25, p.245-281, 2003.

COHEN, Maria Antonieta. *Syntactic change in Portuguese*: relative clauses and the position of the adjective in the Noun Phrase. 1990. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Campinas-SP.

CORBETT, Greville. *Gender*. Cambridge: Cambridge University Press, 1991.

CORNIPS, L.; HULK, A. External and internal factors in bilingual and bidialectal language development: grammatical gender and the Dutch definite determiner. In: LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine (Ed.). *L2 Acquisition and Creole Genesis*: dialogues. Amsterdam: John Benjamins, 2006. p.355-377.

CORRÊA, Letícia Sicuro; AUGUSTO, Marina; FERRARI-NETO, José. The early processing of number agreement in the DP: evidence from the acquisition of Brazilian Portuguese. Trabalho apresentado à 30th Boston University Conference on Language Development, Boston, 2005.

CORRÊA, Vilma R. *Objeto direto nulo no português do Brasil*. 1991. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Campinas-SP.

CORRÊA, Vilma R. *Oração relativa*: o que se fala e o que se aprende no português do Brasil. 1998. Tese (Doutorado) - Universidade Estadual de Campinas, Campinas-SP.

CORRÊA, Vilma R. Aprendendo a relativa padrão na escola. *Cadernos de Estudos Linguísticos*, Campinas, v.36, p.71-83, 1999.

CÔRTES JUNIOR, Moacir J. *Clivadas e pseudoclivadas*: um estudo de suas realizações estruturais no português rural afro-brasileiro. 2006. Dissertação (Mestrado em Linguística) - Universidade Federal da Bahia, Salvador.

COSTA E SILVA, Alberto da. *Um rio chamado Atlântico*: *a África no Brasil e o Brasil na África*. Rio de Janeiro: Editora da UFRJ, 2003.

COSTA, João. Word order and discourse - Configurationality in European Portuguese. In: ______ (Ed.). *Portuguese syntax*: new comparative studies. New York: Oxford University Press, 2000. p.94-115.

COSTA, João. Spec, IP ou deslocado?: prós e contras das duas análises dos sujeitos pré-verbais. *DELTA*, São Paulo, v.17, n.2, p.283-303, 2001. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102445020010002000005&lng=pt&nrm=iso>. Acesso em: 4 mar. 2004.

COSTA, João; GALVES, Charlotte. *Peripheral subjects in two varieties of Portuguese*: evidence for a non-unified analysis. 2000. Disponível em: <http://www.ime.uSP-br/~tycho/papers/ costa_cgalves.pdf>.

COSTA, João; SILVA, Maria Cristina Figueiredo. Nominal and verbal agreement in Portuguese: an argument for distributed morphology. In: COSTA, J.; FIGUEIREDO SILVA, M. C. (Ed.). *Studies on agreement*. Amsterdam: John Benjamins, 2006. p.25-46.

COUTINHO, Ismael de Lima. *Gramática histórica*. 6.ed. Rio de Janeiro: Acadêmica, 1968.

CUNHA, Celso. *Gramática do português contemporâneo*. 9. ed. Rio de Janeiro: Padrão, 1981.

CUNHA, Celso. *Língua portuguesa e realidade brasileira*. Rio de Janeiro: Tempo Brasileiro, 1970.

CUNHA, Celso; CINTRA, Luís F. Lindley. *Nova gramática do português contemporâneo*. 2.ed. Rio de Janeiro: Nova Fronteira, 1985.

CYRINO, Sonia Maria Lazzarini. Observações sobre a mudança diacrônica no português do Brasil: objeto nulo e clíticos. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993. p.163-184.

CYRINO, Sônia Maria Lazzarini. *Objeto nulo no português do Brasil*: um estudo sintático-diacrônico. Londrina: UEL, 1997.

CYRINO, Sônia Maria Lazzarini. O objeto nulo no português do Brasil e no português de Portugal. *Boletim da ABRALIN*, Fortaleza, n.25, p.173-181, 2001.

DAVIDSON, Lisa; GOLDRICK, Matthew A. Tense, agreement and defaults in child Catalan: an optimality theoretic account. In: MONTRUL, Silvina; ORDÓÑEZ, Francisco (Ed.). *Linguistic theory and language development in Hispanic languages*. Sommerville: Cascadilla Press, 2003. p.193-211.

DE CAMP, David. The study of pidgin and creole languages. In: HYMES, D. (Ed.). *Pidginization and creolization of languages*. Cambridge: Cambridge University Press, 1971. p. 13-43.

DECAT, Maria Beatriz Nascimento. Construções de tópico em português: uma abordagem diacrônica à luz do encaixamento no sistema pronominal. In: TARALLO, F. (Org.). *Fotografias sociolinguísticas*. Campinas: Pontes; Editora da UNICAMP, 1989. p.113-137.

DEGRAFF, Michel. *To move or not to move? Placement of verbs and object pronouns in Haitian Creole and in French*. University of Michigan, 1994. Ms.

DEGRAFF, Michel. Creolization, language change, and language acquisition: an epilogue. In: ______ (Ed.). *Language creation and language change*: creolization, diachrony, and development. Cambridge: The MIT Press, 1999. p.473-544.

DEGRAFF, Michel; DEJEAN, Yves. On Haitian Creole's 'very strict' adjacency principle. Trabalho apresentado ao VI Encontro da Society for Pidgin and Creole Linguistics, Boston, 1994.

DEN DIKKEN, Marcel. Specificacional copular sentences and pseudoclefts: a case study. In: EVERAERT, M.; RIEMSDIJK H. van (Ed.). *The blackwell companion to syntax*. v.4. Oxford: Blackweel, 2001. p.5-24.

DETTONI, Rachel do Valle. *A concordância de gênero na anáfora pronominal*: variação e mudança linguística no dialeto da Baixada Cuiabana - Mato Grosso. 2003. Tese (Doutorado em Estudos Linguísticos) - Universidade Federal de Minas Gerais, Belo Horizonte.

DIAS, Juçá F. V. *A concordância de número nos predicativos e particípios passados na fala da Região Sul*. 1996. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Florianópolis.

DUARTE, Maria Eugênia L. *Variação e sintaxe - clítico acusativo, pronome lexical e categoria vazia no português do Brasil.* 1986. Dissertação (Mestrado em Letras) - Pontifícia Universidade Católica de São Paulo, São Paulo.

DUARTE, Maria Eugênia L. Clítico acusativo, pronome lexical e categoria vazia no português do Brasil. In: TARALLO, F. (Org.). *Fotografias sociolinguísticas*. Campinas: Pontes; Editora da Unicamp, 1989. p.19-34.

DUARTE, Maria Eugênia L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993. p.107-128.

DUARTE, Maria Eugênia L. *A perda do princípio "evite pronome" no português brasileiro*. 1995. Tese (Doutorado em Linguística) - Universidade de Campinas, Campinas.

DUARTE, Maria Eugênia L. A sociolinguística paramétrica: perspectivas. In: DA HORA, D.; CHRISTIANO, E. (Org.). *Estudos linguísticos*: realidade brasileira. João Pessoa: Idéia, 1999. p.107-114.

DUPUIS, Fernand. *L'expression du sujet dans les subordonnées en ancien français*. 1989. Tese (Doutorado) - Université de Montréal, Montreal.

ELIA, Sílvio. *A unidade linguística do Brasil*. Rio de Janeiro: Padrão, 1979.

EMBICK, D.; NOYER, R. Movement operations after syntax. *Linguistic Inquiry*, Cambridge, MA: MIT Press, v.32, n.4, p.555-595, 2001.

EMMERICH, Charlotte. O português de contato no Parque Indígena do Xingu. *Estudos Linguísticos e Literários*, Salvador, n.13, p.57-90, jul. 1992.

FARACO, Carlos Alberto. *The imperative sentences in Portuguese*: a semantic and historical discussion. 1982. Tese (Doutorado em Linguística) - University of Salford.

FARACO, Carlos Alberto. *Linguística histórica*: uma introdução ao estudo da história das línguas. São Paulo: Parábola, 2005.

FÁVERO, Leonor Lopes. *Coesão e coerência textuais*. 2.ed. São Paulo: Ática, 1993.

FERRAZ, Luiz Ivens. *The Creole of São Tomé*. Johannesburg: Witwatersrand University Press, 1979.

FERREIRA, Carlota. Remanescentes de um falar crioulo brasileiro. In: FERREIRA, Carlota et al. *Diversidade do português do Brasil*. Salvador: EDUFBA, 1984. p.21-32.

FERREIRA, Graziele. *Cinzento*: memória de uma comunidade negra remanescente de quilombo. 1999. Dissertação (Mestrado) - Pontifícia Universidade de São Paulo, São Paulo.

FIGUEIREDO, Maria Cristina Vieira de. *O objeto direto anafórico no dialeto rural afro-brasileiro*. 2004. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal da Bahia, Salvador.

FRANCESCHINA, Florencia. Case and phi-feature agreement in advanced L2 Spanish grammars. *EUROSLA Yearbook* 2. Amsterdam: John Benjamins, 2002. p.71-86.

FRANCESCHINA, Florencia. *Fossilized second language grammars*: the acquisition of grammatical gender. Amsterdam: John Benjamins, 2005.

FRANCESCHINA, Florencia. Parameterized functional features and SLA. In: LICERAS, Juana M.; ZOBL, Helmut; and GOODLUCK, Helen (Ed.). *Proceedings of the 6th Generative Approaches to Second Language Acquisition*. Somerville: Cascadilla Proceedings Project, 2003. p.97-105.

FRANCO, Paula Vanessa Santos. *Focalização e clivagem*: estudo das construções clivadas no português rural de Portugal. 2007. Dissertação (Mestrado em Linguística) - Universidade Federal da Bahia, Salvador.

FREYRE, Gilberto. *Casa grande e senzala*. 2.ed. Rio de Janeiro: Schmidt, 1936.

FRY, Peter; VOGT, Carlos; GNERRE, Maurício. Mafambura e Caxapura: na encruzilhada da identidade. *Dados - Revista de Ciências Sociais*, Rio de Janeiro, Campus, v.24, n.3, p.373-389, 1981.

FURTADO DA CUNHA, Maria Angélica. Gramaticalização nos mecanismos de negação em Natal. In: MARTELOTTA, Mario Eduardo; VOTRE, Sebastião Josué; CEZARIO, Maria Maura. (Org.). *Gramaticalização no português do Brasil*. Rio de Janeiro: Tempo Brasileiro, 1996. p.167-189.

FURTADO DA CUNHA, Maria Angélica. O modelo das motivações competidoras no domínio funcional da negação. *DELTA*, São Paulo, v.17, n.1, p.1-30, 2001.

GALVES, Charlotte. Pronomes e categorias vazias em português do Brasil. *Cadernos de Estudos Linguísticos*, Campinas, v.7, p.107-136, 1984.

GALVES, Charlotte. O enfraquecimento da concordância no português brasileiro. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993. p.387-408.

GALVES, Charlotte. Tópicos, sujeitos, pronomes e concordância no português brasileiro. *Cadernos de Estudos Linguísticos*, Campinas, v.34, p.19-31, 1998.

GALVES, Charlotte. *Ensaios sobre as gramáticas do português*. Campinas: Editora da UNICAMP, 2001.

GIVÓN, Talmy. Negation. In: _______. *Syntax*: a functional-typological introduction. Amsterdam; Philadelphia: John Benjamins, 1984. v.1.

GOMES, Cristina Abreu. Variação e mudança na expressão do dativo no português brasileiro. In: PAIVA, Maria da Conceição de; DUARTE, Maria Eugênia Lamoglia.

(Org.). *Mudança linguística em tempo real*. Rio de Janeiro: Contracapa; FAPERJ, 2003. v.1, p.81-96.

GONÇALVES, Perpétua. Aspectos da sintaxe do português de Moçambique. In: FARIA, Isabel Hub et al. (Org.). *Introdução à linguística geral e portuguesa*. Lisboa: Caminho, 1996. p.313-322.

GONÇALVES. Perpétua. Towards a unified vision of classes of language acquisition and change: arguments from the genesis of Mozambican African Portuguese. *Journal of Pidgin and Creole Languages*, Amsterdam, v.19, n.2, p.225-259, 2004.

GONÇALVES, Perpétua; STROUD, Christopher. Panorama do português oral de Maputo. v.3: Estruturas gramaticais do português. Maputo: Imprensa Universitária, 1998.

GOODMAN, Morris F. The strange case of Mbugu. In: HYMES, D. (Ed.). *Pidginization and creolization of languages*. Cambridge: Cambridge University Press, 1971. p.243-254.

GOULART, Maurício. *A escravidão africana no Brasil*. 3.ed. São Paulo: Alfa-Ômega, 1975 [1949].

GRACIOSA, Diva. *Concordância verbal na fala culta carioca*. 1991. Dissertação (Mestrado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

GRIMSHAW, Jane. Getting the dative alternation. In: LAKA, Itziar; MAHAJAN, Anoop (Ed.). *The MIT Working Papers in Linguistics*, n.10, p.113-122, 1988.

GUASTI, Maria Teresa; CARDINALETTI, Anna. Relative clause formation in romance child's production. *Probus*, Berlin, n.15, p.47-89, 2003.

GUY, Gregory R. *Linguistic variation in Brazilian Portuguese*: aspects of phonology, sintax and language history. 1981a. Tese (Doutorado em Linguística) - University of Pennsylvania, Pennsylvania.

GUY, Gregory R. Parallel variability in American dialects of Spanish and Portuguese. In: SANKOFF, D.; CEDERGREN, H. (Ed.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, 1981b. p.85-93.

GUY, Gregory R. On the nature and origins of vernacula Brazilian Portuguese. In: *Estudios sobre Español de América y Linguística Afroamericana*. Bogotá: Instituto Caro y Cuervo, 1989. p 226-244.

GUY, Gregory R. A questão da crioulização no português do Brasil. In: ZILLES, Ana Maria Stahl (Org.). *Estudos de variação linguística no Brasil e no Cone Sul*. Porto Alegre: Editora da UFRGS, 2005. p.15-62.

HALLE, Morris. Phonology in a generative grammar. *Word*, Cambridge, n.18, p.54-72, 1962.

HASENBALG, Carlos. *Discriminação e desigualdades raciais no Brasil*. Rio de Janeiro: Graal, 1979.

HAWKINS, Roger. *Second language syntax*: a generative introduction. Oxford: Wiley-Blackwell, 2001.

HAWKINS, Roger; CHAN, Yuet-Hung. The partial availability of universal grammar in second language acquisition: the failed functional features hypothesis. *Second Language Research*, London, v.13, n.3, p.187-226, 1997.

HERNANZ, María Lluisa. L.; BRUCART, José Maria. *La sintaxis. Principios teóricos. La oración simples*. v.1. Barcelona: Crítica, 1987.

HOLM, J. Creole influence on Popular Brazilian Portuguese. In: GILBERT, G. G. (Ed.). *Pidgin and Creole languages*. Essays in Memory of John E. Reinecke. Honolulu: University of Hawaii Press, 1987. p.406-429.

HOLM, John. *Pidgins and Creoles*. Cambridge: Cambridge University Press, 1988.

HOLM, J. Popular Brazilian Portuguese; a semi-creole. In: D'ANDRADE, Ernesto; KIHM, Alain (Org.). *Actas do Colóquio sobre Crioulos de Base Lexical Portuguesa*. Lisboa: Colibri, 1992. p.37-66.

HOLM, John. *Languages in contact*: the partial restructuring of vernaculars. Cambridge: Cambridge University Press, 2004.

HOLM, John. Creolization and the fate of inflections. In: STOLZ, Thomas; BAKKER, Dik; PALOMO, Rosa Salas (Ed.) *Aspects of language contact*: new theoretical, methodological and empirical findings with special focus on romancisation processes. Berlin: Mouton de Gruyter, 2008. p.64-89.

HOLMBERG, Anders. Null subjects and uninterpretable features: evidence from Finnish. *Durham Working Papers in Lingusitics*, Pittsburgh, n.9, p.11-24, 2004.

HOLMBERG, Anders. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry*, Massachusetts, n.36, p.533-564, 2005.

HUANG, James. Pro-Drop in Chinese: a generalized control theory. In: JAEGGLI, O.; SAFIR, K. J. (Ed.). *The null subject parameter*. Dordrecht: Kluwer, 1989. p.185-214.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Estatísticas históricas do Brasil*. v.3. Rio de Janeiro: IBGE, 1987.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Anuário estatístico do Brasil*. v.56. Rio de Janeiro: IBGE, 1996.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. *Contagem da população 1996*. v.1. Rio de Janeiro: IBGE, 1997.

INVERNO, Liliana Cristina Coragem. *Angola's transition to vernacular Portuguese*. 2005. Dissertação (Mestrado em Linguística) - Universidade de Coimbra, Coimbra.

JACKENDOFF, R. *Semantic interpretation in generative grammar.* Cambridge: The MIT Press, 1972.

JESUS, Lavínia Rodrigues de. *O relativo CUJO em documentos do século XIX*. 2002. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal da Bahia, Salvador.

KATO, Mary. Orações relativas: variação universal e variação individual no português. In: SEMINÁRIO DO GEL, *Anais*... Ribeirão Preto, 1981. v.5. p.1-16.

KATO, Mary. A complementaridade dos possessivos e das construções genitivas no português coloquial: réplica a Perini. *DELTA*, São Paulo, v.1, n.1/2, p.107-120, 1985.

KATO, Mary. Tópico e sujeito: duas categorias na sintaxe? In: *Cadernos de Estudos Linguísticos*, Campinas, SP, n.17, p.109-132, 1989.

KATO, Mary. Recontando a história das relativas em uma perspectiva paramétrica. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993. p.223-262.

KATO, Mary. Os frutos de um projeto herético: parâmetros na variação intra-linguística In: DA HORA, D.; CHRISTIANO, E. (Org.). *Estudos linguísticos*: realidade brasileira. João Pessoa: Idéia, 1999a. p.95-106.

KATO, Mary. Strong pronouns, weak pronominals and the null subject parameter. *PROBUS*, Berlin, v.11, n.1, p.1-37, 1999b.

KATO, Mary. A evolução da noção de parâmetros. *DELTA*, São Paulo, v.18, n.2, p.309-338, 2002a.

KATO, Mary. Pronomes fortes e fracos na gramática do português brasileiro. *Revista Portuguesa de Filologia*, Coimbra, v.20, p.101-122, 2002b.

KATO, Mary; MIOTO, Carlos. A multi-evidence study of European and Brazilian Wh-questions. In: KEPSER, S.; REIS, M. (Ed.). *Linguistic evidence*: theoretical, quantitative and computational perspectives. Hague: Mouton, 2005. p.307-330.

KATO, Mary; RAPOSO, Eduardo. European and Brazilian word order: questions, focus and topic constructions. In: PARODI, C. et al. (Ed.). *Aspects of romance linguistics*. Washington: Georgetown U. Press, 1996. p.267-277.

KATO, Mary; RIBEIRO, Ilza. A evolução das estruturas clivadas no português. In: LOBO T. et al. (Org.). *Para a história do português brasileiro*: a experiência dos grupos de estudo. v.6, tomo 1. Salvador: EDUFBA, 2007. p.165-182.

KATO, Mary et al. Construções-Q na gramática do português brasileiro falado: perguntas, clivadas e relativas. In: KOCH, I. G. V. (Ed.). *Gramática do português falado*. v.6: Desenvolvimentos. Campinas: Editora da UNICAMP, 1996. p.303-370.

KAYNE, Richard. Null subjects and clitic climbing. In: JAEGGLI, O.; SAFIR, K. J. (Ed.). *The null subject parameter*. Dordrecht: Kluwer, 1989.

KAYNE, Richard. *Parameters and universals*. Oxford: Oxford University Press, 2000.

KEENAN, Edward L. Relative clauses. In: SHOPEN, Timothy (Ed.). *Language typology and syntactic description*. v.1: Clause structure. Cambridge: Cambridge University Press, 1985. p.141-170.

KEENAN, Edward L.; COMRIE, Bernard. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, n.8, p.63-99, 1977.

KEENAN, Edward L.; COMRIE, Bernard. Data on the noun phrase accessibility hierarchy. *Language*, Baltimore, n.55, p.333-351, 1979.

KIHM, Alain. *Kriyol Syntax*: the Portuguese-Based Creole language of Guinea-Bissau. Amsterdam; Philadelphia: John Benjamins, 1994.

KISS, Katalin. Identificational focus versus information focus. *Language*, n.74, p.245-273, 1998.

KISS, Katalin (Ed.). *Discourse configurational languages*. Oxford: Oxford University Press, 1995.

KLEIN, Wolfgang; PERDUE, Clive. *Utterance structure*: developing grammars again. Amsterdam; Philadelphia: John Benjamins, 1992.

KLEIN, Wolfgang; PERDUE, Clive. The basic variety (or couldn't natural languages be much simpler?). *Second Language Research*, London, v.13, n.4, p.301-347, 1997.

KOEFOED, Geert. Some remarks on the baby talk theory and relexification. In: HANCOCK, I. (Ed.). *Readings in Creole studies*. Ghent: E. Story-Scientia, 1979. p.37-54.

KOOPMAN, Hilda. The genesis of Haitian: implications of a comparison of some features of the syntax of Haitian, French and West African Languages. In: MUYSKEN, Pieter; SMITH, Norval. (Ed.). *Universals versus substrata in Creole genesis*. Amsterdam: John Benjamins, 1986. p.231-259.

KOOPMAN, Hilda; SPORTICHE, Dominique. The position of subject. *Lingua*, Cluj-Napoca, n.85, p.211-258, 1991.

KROCH, Anthony. Reflexes of grammar in patterns of language change. *Language Variation and Change*, Cambridge, n.3, p.199-244, 1989.

KROCH, Anthony. Morphosyntactic variation. In: BEALS, K. et al. (Ed.). *Papers from the 30th Regional Meeting of the Chicago Linguistics Society*. v.2: The Parasession on variation and linguistic theory. Chicago: The University of Chicago Press, v.2. 1994.

KROCH, Anthony. Syntactic change. In: BALTIN, Mark; COLLINS, Chris. (Ed.). *The handbook of contemporary syntactic theory*. Massachusetts: Blackwell , 2001. p.699-729.

KROCH, Anthony. *If at first you don't succeed*: the time course of language acquisition and its implication for language change. 2005. Ms.

KROCH, Anthony; TAYLOR, Anne. Verb movement in Old and Middle English: dialect variation and language contact. In: KEMENADE, A. van; VICENT, N. (Ed.). *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press, 1997. p.297-325.

KROEGER, Paul R. *Analyzing syntax*: a lexical-functional approach. Cambridge: Cambridge University Press, 2004.

LABOV, William. *The social stratification of English in New York city*. Washington: Center of Applied Linguistics, 1966.

LABOV, William. Contraction, deletion and inherent variability of the English copola. *Language*, v.45, p.715-762, 1969.

LABOV, William. *Sociolinguistics patterns*. Oxford: Basil Blackwell, 1972a.

LABOV, William. *Language in the inner city*. Philadelphia: University of Pennsylvania Press, 1972b.

LABOV, William. On the use of the present to explain the past. In: HEILMANN, L. (Ed.). *Proceedings of the Eleventh International Congress of Linguistics*. Bologna: Il Mulino, 1975. p.825-851.

LABOV, William. What can be learned about change in progress from synchrony descriptions. In: SANKOFF, David; CEDERGREN, Henrietta (Ed.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, 1981. p.177-199.

LABOV, William. Building on empirical foundations. In: LEHMANN, W. P.; MALKIEL, Y. (Ed.). *Perspectives on historical linguistics*. Amsterdam; Philadelphia: John Benjamins, 1982. p.17-92.

LABOV, William. *Principles of linguistic change*. Oxford; Cambridge: Blackwell, 1994.

LADGEWAY, Adam. *A comparative syntax of the dialects of Southern Italy*: a minimalist approach. Oxford: Blackwell, 2000.

LAMBRECHT, Knud. A framework for the analysis of cleft constructions. *Linguistics*, v.39, n.3, p.463-516, 2001.

LARSON, Richard. On the double object construction. *Linguistic Inquiry*, 1988. v.1, p.589-632.

LASS, Roger. *On explaining language change*. Cambridge: Cambridge University Press, 1980.

LEFEBVRE, Claire. Variation in plural marking: the case of Cuzco Quechua. In: SANKOFF, D.; CEDERGREN, H. (Ed.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, 1981. p.125-133.

LEFEBVRE, Claire. Creole genesis and the acquisition of grammar: the case of Haitian Creole. Cambridge: Cambridge University Press, 1998.

LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine (Ed.). *L2 acquisition and Creole genesis*: dialogues. Amsterdam: John Benjamins, 2006.

LEMLE, Miriam. Pronomes, anáfora, zero: observações sobre uma mudança linguística. *DELTA*, São Paulo, v.1, n.1/2, p.121-124, 1985.

LEMLE, Mirian; NARO, Anthony. Competências básicas do português. *Relatório final de pesquisa apresentado às instituições patrocinadoras, Fundação MOBRAL e Fundação Ford*. Rio de Janeiro, 1977. Ms.

LEVELT, Willem. *Speaking*: from intention to articulation. Cambridge: MIT Press, 1989.

LI, Charles. Synchrony vs. diachrony in language structure. *Language*, v.51, n.4, p.873-886, 1975.

LIGHTFOOT, David. *Principles of diachronic syntax*. Cambridge: Cambridge University Press, 1979.

LIGHTFOOT, David. *How to set parameters*: arguments from language change. Cambridge: The MIT Press, 1991.

LIGHTFOOT, David. Uma ciência da história? *DELTA*, São Paulo, v.2, n.9, p.275-294, 1993.

LIGHTFOOT, David. Creoles and cues. In: DEGRAFF, Michel (Ed.). *Language creation and language change*: creolization, diachrony, and development. Cambridge: The MIT Press, 1999a. p.431-452.

LIGHTFOOT, David. *The development of language. Acquisition, change, and evolution*. Oxford: Blackwell, 1999b.

LOBATO, Lúcia. A concordância nominal no português do Brasil à luz da Teoria de Princípios e Parâmetros e da Sociolinguística Variacionista. *DELTA*, São Paulo, n.10, p.173-212, 1994. Número especial.

LOPES, Célia Regina dos Santos. Nós e a gente no português falado culto do Brasil, *DELTA*, São Paulo, v.14, n.2, p.405-422, 1998.

LOPES, Célia Regina dos Santos. *A inserção de a gente no quadro pronominal do português*: percurso histórico. 1999. Tese (Doutorado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

LOPES, Nei. *Bantos, malês e identidade negra*. Rio de Janeiro: Forense Universitária, 1988.

LOPES, Norma da Silva. *Concordância nominal, contexto linguístico e sociedade*. 2001. Tese (Doutorado) - Universidade Federal da Bahia, Salvador.

LOPES DA SILVA, Baltasar. O dialeto crioulo de Cabo Verde. Lisboa: Imprensa Nacional-Casa da Moeda, 1984.

LUCCHESI, Dante. The article systems of Cape Verde and São Tomé Creole Portuguese: general principals and specific factors. *Journal of Pidgin and Creole Languages*, Filadélfia; Amsterdam, v.8, n.1, p.81-108, 1993.

LUCCHESI, Dante. Variação e norma: elementos para uma caracterização sociolinguística do português do Brasil. *Revista Internacional de Língua Portuguesa*, Lisboa, n.12, p.17-28, 1994.

LUCCHESI, Dante. Variação, mudança e norma: a questão brasileira. In: CARDOSO, Suzana A. M. (Org.). *Diversidade linguística e ensino*. Salvador: EDUFBA, 1996. p.69-80.

LUCCHESI, Dante. A constituição histórica do português brasileiro como um processo bipolarizado: tendências atuais de mudança nas normas culta e popular. In: GROβE, S.; ZIMMERMANN, K. (Ed.). *"Substandard" e mudança no português do Brasil*. Frankfurt am Main: TFM, 1998. p.73-100.

LUCCHESI, Dante. A questão da formação do português popular do Brasil: notícia de um estudo de caso. *A Cor das Letras*, Feira de Santana, v.3, p.73-100, 1999a.

LUCCHESI, Dante. A variação na concordância de gênero em dialetos despidginizantes e descrioulizantes no português do Brasil. In: ZIMMERMANN, Klaus (Ed.). *Lenguas criollas de base lexical española y portuguesa*. Madri: Iberoamericana, 1999b. p.477-502.

LUCCHESI, Dante. *A variação na concordância de gênero em uma comunidade de fala afro-brasileira*: novos elementos sobre a formação do português popular do Brasil. 2000a. Tese (Doutorado em Linguística) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

LUCCHESI, Dante. O português se teria crioulizado no Brasil? – refletindo sobre uma velha questão. *ABP*, Fankfurt, v.2, p.25-43, 2000b.

LUCCHESI, Dante. As duas grandes vertentes da história sociolinguística do Brasil. *DELTA*, São Paulo, v.17, n.1, p.97-130, 2001a.

LUCCHESI, Dante. O tempo aparente e as variáveis sociais. *Boletim da ABRALIN*, v.26, p.135-137, 2001b. Número especial.

LUCCHESI, Dante. Grandes territórios desconhecidos. *Linguística* (ALFAL), São Paulo, n.14, p.191-222, 2002a.

LUCCHESI, Dante. Norma linguística e realidade social. In: BAGNO, Marcos (Org.). *Linguística da norma*. São Paulo: Loyola, 2002b. p.63-92.

LUCCHESI, Dante. A categoria gramatical do gênero: universais, mudança e crioulização. In: CASTRO, Ivo; DUARTE, Inês. *Razões e emoção*: miscelânea de estudos em homenagem a Maria Helena Mira Mateus. Lisboa: Imprensa Nacional-Casa da Moeda, 2003a. p.429-450.

LUCCHESI, Dante. O conceito de transmissão linguística irregular e o processo de formação do português do Brasil. In: RONCARATI, C.; ABRAÇADO, J. (Org.).

Português brasileiro: contato linguístico, heterogeneidade e história. Rio de Janeiro: 7Letras, 2003b. p.272-284.

LUCCHESI, Dante. A participação do contato entre línguas na formação do português popular do Brasil: novas evidências empíricas. In: M. FERNÁNDEZ, M.; FERNÁNDEZ-FERREIRO, M.; VEIGA, N. V. *Los criollos de base ibérica*. Madrid: Iberoamericana, 2004a. p.215-226.

LUCCHESI, Dante. Contato entre línguas e variação paramétrica: o sujeito nulo no português afro-brasileiro. *Lingua(gem)*, Macapá, v.1, n.2, p.63-92, 2004b.

LUCCHESI, Dante. *Sistema, mudança e linguagem*: um percurso na história da linguística moderna. 2.ed. São Paulo: Parábola, 2004c.

LUCCHESI, Dante. Parâmetros sociolinguísticos do português brasileiro. *Revista da ABRALIN*, v.5, n.1 e 2, p.83-112, 2006a.

LUCCHESI, Dante. Século XVIII, o século da lusofonização do Brasil. In: THIELEMANN, Werner (Ed.). *Século das luzes*: Portugal e Espanha, o Brasil e a Região do Rio da Prata. Frankfurt: TFM, 2006b. p.351-370. (Biblioteca Luso-Brasileira, 24).

LUCCHESI, Dante. Alterações no quadro dos pronomes pessoais e na aplicação da regra de concordância verbal nas normas culta e popular como evidências da polarização sociolinguística do Brasil e da relevância histórica do contato entre línguas. *Linguística* (ALFAL), Santiago, v.19, p.52-87, 2007.

LUCCHESI, Dante. A concordância nominal em estruturas passivas e de predicativo do sujeito em comunidades rurais afro-brasileiras isoladas no contexto da história sociolinguística do Brasil. In: VOTRE, Sebastião; RONCARATI, Cláudia (Org.). *Anthony Julius Naro e a linguística no Brasil*: uma homenagem acadêmica. Rio de Janeiro: 7Letras, 2008a. p.148-168.

LUCCHESI, Dante. Africanos, crioulos e a língua portuguesa. In: LIMA, Ivana Stolze; CARMO, Laura do (Org.). *História social da língua nacional*. Rio de Janeiro: Casa de Rui Barbosa, 2008b. p.151-180.

LUCCHESI, Dante. Aspectos gramaticais do português brasileiro afetados pelo contato entre línguas: uma visão de conjunto. In: RONCARATI, Cláudia; ABRAÇADO, Jussara (Org.). *Português brasileiro II*: contato linguístico, heterogeneidade e história. Niterói: EDUFF, 2008c. p.366-390.

LUCCHESI, Dante; BAXTER, Alan. Processos de crioulização na história sociolinguística do Brasil. In: CARDOSO, Suzana; MOTA, Jacyra; MATTOS E SILVA, Rosa Virgínia (Org.). *Quinhentos anos de história linguística do Brasil*. Salvador: Secretaria da Cultura e Turismo do Estado da Bahia, 2006. p.163-218.

LUMSDEN, John S. Language acquisition and creolization. In: DEGRAFF, Michel (Ed.). *Language creation and language change*: creolization, diachrony, and development. Cambridge: The MIT Press, 1999a. p.129-157.

LUMSDEN, John. S. The role of relexification in creole genesis. *Journal of Pidgin and Creole Languages*. Amsterdam, v.14, n.2, p.225-258, 1999b.

LUMWANU, François. *Essai de morphosyntaxe systématique des parlers kongo*. Paris: Klincksieck, 1973.

LYONS, John. *Introdução à linguística teórica*. Trad. Rosa Virgínia Mattos e Silva e Hélio Pimentel, rev. e superv. de Issac N. Salum. São Paulo: Nacional, 1979.

LYONS, John. *Definiteness*. Cambridge: Cambridge University Press, 1999.

MACHADO, Márcia. Sujeitos pronominais nós e a gente em dialetos populares. *Graphos*, v.2, n.1, p.5-24, 1997.

MACHADO FILHO, Aires da M. *O negro e o garimpo em Minas Gerais*. São Paulo: Itatiaia, 1985.

MALER, Bertil. *O orto do esposo*. Glossário. Stokholm: Almqvist & Wiksell, 1964.

MARQUILHAS, Rita. *Leitura e escrita em Portugal no século XVII*. 1997. Tese (Doutorado em Linguística) - Universidade de Lisboa, Lisboa.

MARTINEZ, Maria Sol López. *O complemento direto com preposición "A" em galego*. Santiago de Compostela: Universidade de Santiago de Compostela, 1993.

MASCARENHAS, Salvador. *Complementizer doubling in European Portuguese*. 2007. Ms.

MATA, Ariadne Costa da. A variável da negação na comunidade de João Pessoa. In: SOARES, Maria Elias; ARAGÃO, Maria do Socorro Silva de. (Org.). In: JORNADA DE ESTUDOS LINGUÍSTICOS DO NORDESTE, 16., 1998, Fortaleza. *Anais*... Fortaleza: UFC; GELNE, 1999. v.1. p.344-347.

MATEUS, Maria Helena Mira et al. *Gramática da língua portuguesa*. 5.ed. Lisboa: Caminho, 2003.

MATHER, Patrick-André. Creole genesis: evidence from West African L2 French. In: GILBERS, D. G., NERBONNE, J.; SCHAEKEN, J. (Ed.). *Languages in Contact*. Amsterdam; Atlanta: Rodopi, 2000. p.247-261.

MATHER, Patrick-André. Second language acquisition and creolization: Same (i-) processes, different (e-) results. *Journal of Pidgin and Creole Languages*, Amsterdam, v.21, n.2, p.231-274, 2006.

MATTOS E SILVA, Rosa Virgínia. (Org.). *Sete estudos sobre o português kamayurá*. Salvador: UFBA; PROED, 1988.

MATTOS E SILVA, Rosa Virgínia. *Estruturas trecentistas*: elementos para uma gramática do português arcaico. Lisboa: Imprensa Nacional-Casa da Moeda, 1989.

MATTOS e SILVA, Rosa Virgínia. Idéias para a história do português brasileiro: fragmentos para uma composição posterior. In: CASTILHO, Ataliba de (Org.). *Para a história do português brasileiro*. v.1: Primeiras idéias. São Paulo: Humanitas, 1998. p.21-52.

MATTOS E SILVA, Rosa Virgínia. *Ensaios para uma sócio-história do português brasileiro*. São Paulo: Parábola, 2004.

MATTOSO, Katia. *Ser escravo no Brasil*. 3.ed. São Paulo: Brasiliense, 2003.

McCARTHY, Corrine. Default morphology in second language Spanish: missing inflection or underspecified inflection? In: NISHIDA, Chiyo; MONTREUIL, Jean-Pierre Y. (Ed.). *New perspectives on romance linguistics*: selected papers from the 35th Linguistic Symposium on Romance Languages. Amsterdam: John Benjamins, 2006. p.201-212.

MEGENNEY, William W. Helvecian Portuguese: vernacular dialect or true creole? *Journal of Pidgin and Creole Languages*, Columbus-Ohio, v.2, p.237-240, 1993.

MEINTEL, Deirdre. The Creole dialect of tht Island of Brava. In: VALKHOFF, Marius F. (Org.). *Miscelânea Luso-Africana*. Lisboa: Junta de Investigações do Ultramar, 1975. p.205-256.

MEIRA, Vivian. *O uso do modo subjuntivo em orações relativas e completivas no português afro-brasileiro*. 2006. Dissertação (Mestrado em Letras e Linguística) - Universidade Federal da Bahia, Salvador.

MELO, Gladstone Chaves de. *A língua do Brasil*. Rio de Janeiro: Agir, 1946.

MENDES, Beatriz C. *Contributo para o estudo da língua portuguesa em Angola*. Lisboa: Instituto de Linguística da Faculdade de Letras da Universidade de Lisboa, 1985.

MENDONÇA, Renato. *A influência africana no português do Brasil*. Rio de Janeiro: Sauer, 1933.

MINGAS, Amélia A. *Interferência do kimbundu no português falado em Luanda*. Lisboa: Campo das Letras, 2000.

MODESTO, Marcelo. *As construções clivadas no português do Brasil*. São Paulo: Humanitas, 2001.

MONGUILHOTT, Isabel. Variação na concordância verbal de 3ª pessoa do plural na fala dos florianopolitanos. 2001. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Florianópolis.

MONGUILHOTT, Isabel; COELHO, Izete. Um estudo da concordância verbal de terceira pessoa em Florianópolis. In: VANDRESEN, Paulino (Org.). *Variação e mudança no português falado na Região Sul*. Pelotas: EDUCAT, 2002. p.189-216.

MONTEIRO, José Lemos. *Os pronomes pessoais no português do Brasil*. 1991. Tese (Doutorado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

MONTEIRO, José Lemos. *Pronomes pessoais*: subsídios para uma gramática do português do Brasil. Fortaleza: Edições UFC, 1994.

MONTRUL, Silvina A. Incomplete acquisition in bilingualism as an instance of language change. In: LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine (Ed.). *L2 acquisition and Creole genesis*: dialogues. Amsterdam: John Benjamins, 2006. p.379-400.

MORAIS-BARBOSA, Jorge. Cape Verde, Guinea-Bissau and São Tomé and Príncipe: the linguistic situation. In: VALKHOFF, Marius F. (Org.). *Miscelânea Luso-Africana*. Lisboa: Junta de Investigações do Ultramar, 1975. p.133-151.

MORENO CABRERA, Juan Carlos. Las funciones informativas: las perífrasis de relativo y otras construcciones perifrásticas. In: BOSQUE, I.; DEMONTE, V. (Org.). *Gramática descriptiva de la lengua española*. v.3. Madri: Espasa Calpe, 1999. p.4245-4302.

MUFWENE, S. S. Transfer and the substrate hypothesis in creolistics. *Studies in Second Language Acquisition*, Cambridge, v.12, n.1, p.1-23, 1990.

MUFWENE, Salikoko. Grammaticization is part of the development of creoles. *Papia*, Brasília, n.16, p.5-31, 2006.

MÜHLHÄUSLER, Peter. *Pidgin & Creole linguistics*. Londres: Basil Blackwell, 1986.

MÜLLER, Ana Lúcia. *A gramática das formas possessivas no português do Brasil*. 1997. Tese (Doutorado em Linguística) - Universidade Estadual de Campinas, Campinas-SP.

MUSSA, Alberto B. N. *O papel das línguas africanas na história do português do Brasil*. 1991. Dissertação (Mestrado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

MUYSKEN, Pieter; VEENSTRA, Tonjes. Universalist approaches. In: ARENDS, Jacques; MUYSKEN, Pieter; NORVAL, Smith. *Pidgins and creoles*: an introduction. Amsterdam; Philadelphia: John Benjamis, 1994. p.121-136.

MYERS-SCOTTON, Carol. *Duelling languages*: grammatical structure in codeswitching. New York: Oxford University Press, 1997.

MYERS-SCOTTON, Carol; JAKE, Janice L. Four types of morpheme: evidence from aphasia, code switching, and second-language acquisition. *Linguistics*: *An Interdisciplinary Journal of the Language Sciences*, Berlin, v.38, n.6, p.1053-1100, 2000.

NARO, Anthony. The social and structural dimensions of a syntactic change. *Language*, v.57, n.1, p.63-98, 1981.

NARO, Anthony. O dinamismo das línguas. In: MOLLICA, Cecília; BRAGA, Maria Luiza (Org.). *Introdução à sociolinguística*. São Paulo: Contexto, 2003a. p.43-50.

NARO, Anthony. Modelos quantitativos e tratamento estatístico. In: MOLLICA, Cecília; BRAGA, Maria Luiza (Org.). *Introdução à sociolinguística*. São Paulo: Contexto, 2003b. p.5-26.

NARO, Anthony; GÖRSKI, E.; FERNANDES, E. Change without change. *Language Variation and Change*, New York, v.11, n.2, p.197-211, 1999.

NARO, Anthony; LEMLE, Miriam. Syntatic diffusion. *Ciência e Cultura*, v.29, n.3, p.259-268, 1976.

NARO, Anthony; SCHERRE, Marta. Variação e mudança linguística: fluxos e contrafluxos na comunidade de fala. *Cadernos de Estudos Linguísticos*, Campinas, n.20, p.9-16, 1991.

NARO, Anthony; SCHERRE, Marta. Sobre as origens do português popular do Brasil. *DELTA*, São Paulo, v.9, p.437-454, 1993. Número especial.

NARO, Anthony; SCHERRE, Marta. A influência de variáveis escalares na concordância verbal. *A Cor das Letras*, Feira de Santana, n.3, p.17-34, 1999.

NARO, Anthony; SCHERRE, Marta. Variable concord in Portuguese: the situation in Brazil and Portugal. In: McWHORTER, J. (Ed.). *Language change and language contact in pidgins and creoles*. Amsterdam: John Benjamins, 2000. p.235-256.

NARO, Anthony; SCHERRE, Marta. *Origens do português brasileiro*. São Paulo: Parábola, 2007.

NEVES, Maria Helena de Moura. Possessivos. In: CASTILHO, Ataliba Teixeira de (Org.). *Gramática do português falado*. v.3: As abordagens. Campinas: Editora da Unicamp; FAPESP, 1993. p.149-213.

NEVES, Maria Helena de Moura. *Gramática de usos do português*. São Paulo: Editora da UNESP, 2000.

NINA, Terezinha. *Concordância nominal/verbal do analfabeto na Micro-Região Bragantina*. 1980. Dissertação (Mestrado em Letras) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.

NUNES, J. Joaquim. *Compêndio de gramática histórica portuguesa*. 6.ed. Lisboa: Clássica, 1960.

OLIVEIRA, Klebson. *Negros e escrita no Brasil do século XIX*: sócio-história, edição filológica de documentos e estudo linguístico. 2006. Tese (Doutorado em Letras e Linguística) - Universidade Federal da Bahia, Salvador.

OMENA, Nelize. *Pronome pessoal de terceira pessoa*: suas formas variantes em função acusativa. 1978. Dissertação (Mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro.

OMENA, Nelize. A referência à primeira pessoa do discurso no plural. In: OLIVEIRA, Giselle Machline de; SCHERRE, Maria Marta Pereira (Org.). *Padrões sociolinguísticos*: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996a. p.183-215.

OMENA, Nelise. As influências sociais na variação *nós* e *a gente* na função de sujeito. In: OLIVEIRA, Giselle Machline de; SCHERRE, Marta (Org.). *Padrões sociolinguísticos*: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro. Rio de Janeiro: Tempo Brasileiro, 1996b. p.309-324.

OMENA, Nelise. A referência à primeira pessoa do plural: variação ou mudança? In: PAIVA, Maria da Conceição de; DUARTE, Maria Eugênia Lamoglia. (Org.). *Mudança linguística em tempo real*. Rio de Janeiro: Contracapa; FAPERJ, 2003. v.1. p.63-80.

ORDÓÑEZ, Salvador Gutiérrez. Los dativos. In: BOSQUE, I; DEMONTE, V. (Org.). *Gramática descriptiva de la lengua española*. Madrid: Espasa-Calpe, 1999. p.1855-1930.

PAGOTTO, Emilio Gozze. Norma e condescendência: ciência e pureza. *Línguas e Instrumentos Linguísticos*, Belo Horizonte, v.2, 1998. p.49-68.

PEDREIRA, Pedro Tomás. Os quilombos baianos. In: MOURA, Clóvis (Org.). *Os quilombos na dinâmica social do Brasil*. Maceió: EDUFAL, 2001. p.233-254.

PEREIRA, Maria Ângela Botelho. *Gênero e número em português*. Rio de Janeiro: PROED/UFRJ, 1987.

PERINI, Mário. O surgimento do sistema possessivo do português coloquial: uma abordagem funcional. *DELTA*, São Paulo, v.1, n.1/2, p.1-16, 1985.

PERRONI, Cecília. As relativas que são fáceis na aquisição do português brasileiro. *DELTA*, São Paulo, v.17, n.1, p.59-79, 2001.

PICALLO, C. Nominals and nominalization in Catalan. *Probus*: *International Journal of Latin and Romance Linguistics*, Berlin, v.3, n.3, p.279-316, 1991.

PIERSON, Donald. *Brancos e pretos na Bahia*. São Paulo: Nacional, 1945.

PIMPÃO, Tatiana. *Variação no presente do modo subjuntivo*: uma abordagem discursivo-pragmática. 1999. Dissertação (Mestrado em Letras) - Universidade Federal de Santa Catarina, Florianópolis.

PINTZUK, Susan. *VARBRUL programs*. Inédito. Tradução de Ivone Isidoro Pinto, revisão de Maria Thereza Gomes Fioreti e coordenação de Maria Marta Pereira Scherre. 1988. Ms.

PONTES, Eunice S. L. *Sujeito*: da sintaxe ao discurso. São Paulo: Ática, 1986.

PONTES, Eunice S. L. *O tópico no português do Brasil*. Campinas: Pontes, 1987.

POPLACK, Shana. The notion of the plural in Puerto Rican Spanish: competing constraints on /s/ deletion. In: LABOV, W. (Ed.). *Locating language in time and space*. Philadelphia: University of Pennsylvania, 1979. p.55-67.

POPLACK, Shana. Deletion and disambiguation in Puerto Rican Spanish. *Language*, v.56, n. 2, p.371-385, 1980a.

POPLACK, Shana. Sometimes I'll start a sentence in Spanish y termino en espanol: toward a typology of code-switching. *Linguistics*, v.18, n.7/8, p.581-618, 1980b.

POPLACK, Shana. Mortal phonemes as plural morphemes. In: SANKOFF, D.; CEDERGREN, H. (Ed.). *Variation Omnibus*. Carbondale; Edmonton: Linguistic Research, 1981. p.59-71.

PRADO JR., Caio. *História econômica do Brasil*. 17.ed. São Paulo: Brasiliense, 1974.

QUINT, Nicolas. A realização do sujeito no português do Brasil: deriva *versus* crioulização. In: FIORIN, José Luiz; PETTER, Margarida. *África no Brasil*: a formação da língua portuguesa. São Paulo: Contexto, 2008. p.75-88.

RADFORD, Andrew. *Transformational grammar*: a first course. Cambridge: Cambridge University Press, 1988.

RADFORD, Andrew; PLOENNING-PACHECO, Ingrid. The morphosyntax of subjects and verbs in child Spanish. *Essex Research reports in Linguistics*, Colchester, v.5, p.23-67, 1995.

RAIMUNDO, Jacques. *O elemento afro-negro na língua portuguesa*. Rio de Janeiro: Renascença, 1933.

RAMOS, Jânia. "Sociolinguística paramétrica" ou "Variação paramétrica". In: DA HORA, D.; CHRISTIANO, E. (Org.). *Estudos linguísticos*: realidade brasileira. João Pessoa: Idéia, 1999. p.83-93.

RAMOS, Jânia. A alternância "não" e "num" no dialeto mineiro: um caso de mudança linguística. In: COHEN, Maria Antonieta A. M.; RAMOS, Jânia M. (Org.). *Dialeto mineiro e outras falas*. Belo Horizonte: Faculdade de Letras UFMG, 2002. v.1, p.155-167.

RAPOSO, Eduardo. On the null object in European Portuguese. In: JAEGGLI, O.; SILVA-CORVALÁN, C. (Ed.). *Studies in romance linguistics*. Foris: Dordrecht, 1986. p.373-390.

RAPOSO, Eduardo. *Towards a unification of topic constructions*. UCSB, 1996. Inédito.

RAPOSO, Eduardo. Objectos nulos e CLLD: uma teoria unificada. *Revista da ABRALIN*, Maceió, v.3, n.1, p.41-73, 2004.

RIBEIRO, Darcy. *O povo brasileiro*: evolução e sentido do Brasil. 2.ed. 10.reimp. São Paulo: Companhia das Letras, 1997.

RIBEIRO, Ilza. *O efeito V2 no português arcaico*. 1995. Tese (Doutorado em Lingüística) - Universidade de Campinas, Campinas-SP.

RIBEIRO, Ilza. Aplicação de princípios e parâmetros da teoria da gramática ao estudo do fenômeno de *code-switching*. *DELTA*, São Paulo, v.12, n.2, p.327-366, 1996.

RIBEIRO, Ilza. As mudanças sintáticas do PE - questões sobre periodização. In: CASTILHO, A. T. de et al. (Org.). *Descrição, história e aquisição do português brasileiro*. São Paulo: Fapesp; Campinas: Pontes, 2007. p.529-547.

RIBEIRO, Ilza. O sujeito nulo referencial no português popular brasileiro - século XIX. In: MORAIS, M. Aparecida Torres; ANDRADE, M. Lucia O. (Org.). *História do português paulista*. Campinas: Editora da UNICAMP. No prelo.

RIBEIRO, Ilza; FIGEIREDO, Cristina. *As sentenças relativas em Atas escritas por africanos, no Brasil oitocentista (1832-1842)*. Salvador, 2006. Ms.

RIBEIRO, Maria de L. B. O jongo. *Revista do Arquivo Municipal,* São Paulo, n.173, p.165-238, 1968.

RICKFORD, John. *Dimensions of a creole continuum*: history, texts and linguistic analysis of Guyanese Creole. Stranford: Stranford University Press, 1987.

RISÉRIO, Antonio. *Uma história da Cidade da Bahia*. Rio de Janeiro: Versal, 2004.

RIZZI, Luigi. *Issues in Italian syntax*. Dordrecht: Foris, 1982.

RIZZI, Luigi. The fine structure of the left periphery. In: HAEGEMAN, Liliane. (Ed.). *Elements of grammar*: handbook of generative syntax. Dordrecht: Kluwer, 1997. p.281-337.

RIZZI, Luigi. *On the position "int(errogative)" on the left periphery of the clause*. 1999. Disponível em: <http://www.ciscl.unisi.it/doc/doc_pub/int.doc>. Acesso em: 1999.

RIZZI, Luigi. *On the form of chains*: criterial positions and ECP effects. 2003. Disponível em: <http://www.ciscl.unisi.it/doc/doc-pub/rizzi>. Acesso em: 2004.

ROBERTS, Ian. *Verbs and diachronic syntax*: a compative history of English and French. Dordrecht: Kluwer, 1993.

ROBERTS, Ian. Creoles, markedness and the Language Bioprogram Hypothesis. *Estudos Linguísticos e Literários*, Salvador, n.19, p.11-24, 1997.

ROBERTS, Ian. Verb movement and markedness. In: DEGRAFF, Michel (Ed.). *Language creation and language change*: creolization, diachrony, and development. Cambridge: The MIT Press, 1999. p.287-328.

ROBERTS, Ian. *Diacronic syntax*. Oxford: Oxford University Press, 2007.

ROBERTS, Sarah J. *The emergence of Hawai'i Creole English in the early 20th century*: the sociohistorical context of Creole genesis. 2005. Tese (Doutorado) - Stanford University.

ROCHA LIMA, Carlos Henrique da. *Gramática normativa da língua portuguesa*. 5.ed. Rio de Janeiro: F. Briguiet, 1960.

ROCHA, Rosa C. F. *A alternância indicativo/subjuntivo nas orações subordinadas substantivas em português*. 1997. Dissertação (Mestrado em Linguística) - Universidade de Brasília, Brasília.

RODRIGUES, Ada Natal. *O dialeto caipira na região de Piracicaba*. São Paulo: Ática, 1974.

RODRIGUES, Ângela. *Língua e contexto sociolinguístico*: concordância no português popular de São Paulo. Araraquara: UNESP, 1992.

RODRIGUES, Aryon D. *Línguas brasileiras*: para o conhecimento das línguas indígenas. São Paulo: Loyola, 1986.

RODRIGUES, José Honório. A vitória da língua portuguesa no Brasil Colonial. *Humanidades*, Brasília, v.1, n.4, p.21-41, 1983.

RODRIGUES, Nina. *Os africanos no Brasil*. 8.ed. Brasília: Editora Universidade de Brasília, 2004 [1993].

ROMAINE, Suzanne. *Socio-historical linguistics*: its status and methodology. Cambridge: Cambridge University Press, 1982.

ROMAINE, Suzanne. *Pidgin and creole languages*. New York: Longman, 1988.

ROMAINE, Suzanne. Variation. In: DOUGHTY, Catherine J.; LONG, Michael L. (Ed.). *The handbook of second language acquisition*. Oxford: Blackwell, 2003. p.409-434.

RONCARATI, Cláudia. A negação no português falado. In: MACEDO, Alzira Tavares de; RONCARATI, Cláudia; MOLLICA, Maria Cecília. (Org.). *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro, 1996. p.97-112.

ROUGÉ, Jean-Louis. Les langues des Tonga. In: D'ANDRADE, Ernesto; KIHM, Alain (Org.). *Actas do Colóquio sobre Crioulos de Base Lexical Portuguesa*. Lisboa: Colibri, 1992. p.171-176.

ROUGÉ, Jean-Louis. A inexistência de crioulo no Brasil. In: FIORIN, José Luiz; PETTER, Margarida. *África no Brasil*: a formação da língua portuguesa. São Paulo: Contexto, 2008. p.63-74.

ROUSSEAU, Pascale; SANKOFF, David. Advances in variable rule methodology. In: SANKOFF, David (Ed.). *Linguistics variation*: models and methods. New York: Academic Press, 1978. p.57-68.

SAID ALI, Manoel. *Gramática secundária e gramática histórica da língua portuguesa*. 3.ed. Brasília: Editora Universidade de Brasília, 1964.

SAID ALI, Manoel. *Gramática histórica da língua portuguesa*. São Paulo: Melhoramentos; Brasília: Editora Universidade de Brasília, 2001 [1921].

SÁNCHEZ, Liliana. Bilingual grammars and Creoles: similarities between functional convergence and morphological elaboration. In: LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine (Ed.). *L2 acquisition and Creole Genesis*: dialogues. Amsterdam: John Benjamins, 2006.p.277-294.

SANKOFF, David. Variable rules. In: AMMON, Ulrich; DITTMAR, Norbert; MATTHEIR, Klauss (Ed.). *Sociolinguistics*: an international handbook of the science of language and society. Berlin; New York: Walter de Guyter, 1988. p.984-998.

SANKOFF, David; TAGLIAMONTE, Sali; SMITH, Eric. *Goldvarb X*: a variable rule application for Macintosh and Windows. Department of Linguistics, University of Toronto, 2005.

SANKOFF, Gillian. Substrate and universals in the Tok Pisin verb phrase. In: SCHIFFRIN, Deborah (Ed.). *Meaning, form and use in context*: linguistic applications. Washington, DC: Georgetown University Press, 1984. p.104-119.

SANKOFF, Gillian. An historical approach to variation in the Tok Pisin verb phrase. In: BEALS, K. et al. (Ed.). *Papers from the 30th Regional Meeting of the Chicago Linguistics Society*. v.2: The Parasession on variation and linguistic theory. Chicago: The University of Chicago Press, 1994. v.2, p.293-320.

SANTOS, Maria Joana de Almeida Vieira dos. *Os usos do conjuntivo em língua portuguesa*: uma proposta de análise sintáctica e semântico-pragmática. Lisboa: Fundação Calouste Gulbenkian, 2003.

SCHER, Ana Paula. *As construções com dois complementos no inglês e no português do Brasil*: *um estudo sintático comparativo*. 1996. Dissertação (Mestrado em Linguística) - Universidade Estadual de Campinas, Campinas-SP.

SCHERRE, Maria Marta Pereira. *Reanálise da concordância nominal em português*. 1988. Tese (Doutorado em Letras) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

SCHERRE, Marta. A concordância de número nos predicativos e particípios passados. *Organon*, Porto Alegre, v.18, p.52-70, 1991.

SCHERRE, Marta. Aspectos da concordância de número no português do Brasil. *Revista Internacional de Língua Portuguesa*, Lisboa, n.12, p.37-49, 1994.

SCHERRE, Marta. Pressupostos teóricos e suporte quantitativo. In: OLIVEIRA, Giselle Machline de; SCHERRE, Marta (Org.). *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996a. p.37-50.

SCHERRE, Marta. Sobre a influência de três variáveis relacionadas na concordância nominal em português. In: OLIVEIRA, Giselle Machline de; SCHERRE, Marta (Org.). *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996b. p.85-118.

SCHERRE, Marta. P. Phrase-level parallelism effect on noun phrase number agreement. *Language Variation and Change*, Cambridge, v.13, n.1, p.91-107, 2001.

SCHERRE, Marta; NARO, Anthony. Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil. *DELTA*, São Paulo, v.9, n.1, p.1-14, 1993.

SCHERRE, Marta; NARO, Anthony. A concordância de número no português do Brasil: um caso típico de variação inerente. In: HORA, Dermeval da (Org.). *Diversidade linguística no Brasil*. João Pessoa: Idéia, 1997. p.93-114.

SCHERRE, Marta; NARO, Anthony. Análise quantitativa e tópicos de interpretação do VARBRUL. In: MOLLICA, Cecília; BRAGA, Maria Luiza (Org.). *Introdução à sociolinguística*. São Paulo: Contexto, 2003. p.147-178.

SCHIFFRIN, Deborah. Tense variation in narrative. *Language*, v.57, n.1, p.5-62, 1981.

SCHUMANN, John H. *The pidginization process*: a model for second language acquisition. Rowley, MA: Newberry House, 1978.

SCHWARTZ, Bonnie D.; SPROUSE, Rex. L2 cognitive states and the full transfer/full access model. *Second Language Research*, London, v.12, n.1, p.40-72, 1994.

SCHWEGLER, Armin. Palenquero. In: PERL, Matthias; SCHWEGLER, Armin (Ed.). *América negra*: panorámica actual de los estudios lingüísticos sobre variedades criollas y afrohispanas. Frankfurt: Vervuert, 1998. p.220-291.

SCHWEGLER, Armin. Bare nouns in Palenquero: a fresh consensus in the making. In: BAPTISTA, Marlyse; GUÉRON, Jacqueline (Ed.). *Noun phrases in creole languages*: a multi-faceted approach. Amsterdam; Philadelphia: John Benjamins, 2007. p.12-34.

SEARA, Izabel. A variação do sujeito nós e a gente na fala florianopolitana. *Organon*, Porto Alegre, v.14, n.28-29, p.179-94, 2000.

SERRANO, Jônatas. *História do Brasil*. 2.ed. Rio de Janeiro: F. Briguiet, 1968.

SEUREN, Pieter A.M.; WEKKER, Herman C. Semantic transparency as factor in Creole genesis. In: MUYSKEN, P.; SMITH, N. (Ed.). *Substrata versus universals in Creole genesis*. Amsterdam: John Benjamins, 1986. p.57-70.

SIEGEL, Jeff. Transfer constraints and substrate influence in Melanesian pidgin. *Journal of Pidgin and Creole Languages*, Amsterdam, v.14, n.1, p.1-44, 1999.

SIEGEL, Jeff. Substrate influence in Hawaií Creole English. *Language in Society*, Cambridge, v.29, p.197-236, 2000.

SIEGEL, Jeff. Substrate influence in creoles and the role of transfer in second language acquisition. *Studies in Second Language Acquisition*, Cambridge, v.25, n.2, p.185-209, 2003.

SIEGEL, Jeff. Morphological simplicity in pidgins and creoles. *Journal of Pidgin and Creole Languages*, Amsterdam, v.19, n.1, p.139-162, 2004.

SIEGEL, Jeff. Links between SLA and Creole studies: past and present. In: LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine (Ed.). *L2 acquisition and Creole Genesis*: dialogues. Amsterdam: John Benjamins, 2006. p.15-46.

SILVA, Giselle Machline de Oliveira e. *Estudo da regularidade na variação dos possessivos no português do Rio de Janeiro*. 1982. Tese (Doutorado em Língua Portuguesa) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

SILVA, Giselle Machline de Oliveira e. Variação no sistema possessivo de terceira pessoa. *Tempo Brasileiro*, Rio de Janeiro, n.78/79, p.54-72, 1984.

SILVA, Giselle Machline de Oliveira e. Estertores da forma seu na língua oral. In: SILVA, Giselle Machline de Oliveira e; SHERRE, Maria Marta (Org.). *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996. p.171-182.

SILVA, Giselle Machline Oliveira e. Realização facultativa do artigo definido diante de possessivo e de patronímico. In: SILVA, Giselle Machline de Oliveira e; SCHERRE, Maria Marta (Org.). *Padrões sociolinguísticos*. Rio de Janeiro: Tempo Brasileiro, 1996. p.119-145.

SILVA, Jorge Augusto Alves da. *A concordância verbal no português afro-brasileiro*: um estudo sociolinguístico de três comunidades rurais do Estado da Bahia. 2003. Dissertação (Mestrado em Linguística) - Universidade Federal da Bahia, Salvador.

SILVA-CORVALÁN, Carmen. (Ed.) *Studies in Romance Linguistics*. Foris: Dordrecht, 1986.

SILVA NETO, Serafim da. *Introdução ao estudo da língua portuguesa no Brasil*. 2.ed. Rio de Janeiro: INL, 1963 [1951].

SIMÕES, Maria C. P.; STOEL-GAMMON, Carol. The acquisition of inflections in Portuguese: a study of the development of person markers on verbs. *Journal of Child Language*, n.6, p.53-67, 1979.

SIMONSEN, Roberto. *História econômica do Brasil (1500-1820)*. 2v. São Paulo: Nacional, 1937.

SLOBIN, Dan I. The acquisition and use of relative clauses in Turkic and Indo-European languages. In: SLOBIN, D. I.; ZIMMER, K. (Ed.). *Studies in Turkish linguistics*. Amsterdam: John Benjamins, 1986. p.1-15.

SMITH, Ian R. Siri Lanka Creole Portuguese phonology. *International Journal of Davidian Linguistics*, n.7, p.247-406. 1977.

SMITH, Norval. Very rapid Creolization in the framework of the restricted motivation hypothesis. In: LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine

(Ed.). *L2 acquisition and Creole Genesis*: dialogues. Amsterdam: John Benjamins, 2006. p.49-65.

SOUZA, Arivaldo Sacramento de. *As estruturas de negação em uma comunidade rural afro-brasileira*: Helvécia-BA, 2004. Disponível em <http://www.vertentes.ufba.br/souza.doc>. Acesso em: 18 set. 2006.

SOUZA ANDRADE, M. J. de. *A mão de obra escrava em Salvador, 1811-1860*. São Paulo: Corrupio, 1988.

SPROUSE, Rex. Full Transfer and Relexification: Second Language Acquisition and Creole Genesis. In: LEFEBVRE, Claire; WHITE, Lydia; JOURDAN, Christine (Ed.). *L2 acquisition and Creole Genesis*: dialogues. Amsterdam: John Benjamins, 2006. p.169-181.

SUÑER, Margarita. Resumptive restrictive relatives: a crosslinguistic perspective. *Language*, Baltimore, n.74, p.335-364, 1998.

TAGLIAMONTE, Sali A. *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press, 2006.

TARALLO, Fernando. *Relativization strategies in Brazilian Portuguese*. 1983. Tese (Doutorado) - University of Pennsylvania, Pennsylvania.

TARALLO, Fernando. Por uma sociolinguística românica paramétrica: fonologia e sintaxe. *Cadernos de Linguística e Teoria da Literatura*, Belo Horizonte, v.7, n.13, p.51-83, 1987.

TARALLO, Fernando. Reflexões sobre o conceito de mudança linguística. *Organon*, Porto Alegre, v.1, n.1, p.11-22, 1991.

TARALLO, Fernando. Diagnosticando uma gramática brasileira: o português d'aquém e d'além-mar ao final do século XIX. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993a. p.9-106.

TARALLO, Fernando. Sobre a alegada origem crioula do português brasileiro: mudanças sintáticas aleatórias. In: KATO, Mary A.; ROBERTS, Ian (Org.). *Português brasileiro*: uma viagem diacrônica. Campinas: Editora da Unicamp, 1993b. p.35-68.

TARALLO, Fernando; KATO, Mary A. Harmonia trans-sistêmica: variação inter e intra-linguística. *Preedição*, Campinas, n.6, p.1-41, 1989.

TORRES MORAIS, Maria Aparecida; RIBEIRO, Ilza; FERREIRA, Michael J. Rastreando dativos de 3ª. pessoa na *Grammatica* de Fernão de Oliveira. No prelo.

VANCE, B. S. *Null subjects and syntactic change in Medieval French*. Ann Arbor: University Microfilms, 1989.

VAREJÃO, Filomena de Oliveira Azevedo. *Variação em estruturas de concordância verbal e em estratégias de relativização no português europeu popular*. 2006. Tese (Doutorado) - Universidade Federal do Rio de Janeiro, Rio de Janeiro.

VASCONCELOS, Carolina Michaëlis de. *Cacioneiro da Ajuda*. Reimpressão da edição Halle (1904), acrescentada de um prefácio de Ivo Castro e do glossário das cantigas (Revista Lusitana, v.23). Lisboa: Imprensa Nacional-Casa da Moeda, 1990.

VEADO, Rosa Maria A. *Comportamento linguístico do dialeto rural*. Belo Horizonte: UFMG; PROED, 1982.

VIEIRA, Sílvia. A não concordância em dialetos populares: uma regra variável. *Grafos,* João Pessoa, v.2, n.1, p.115-134, 1997.

VOGT, Carlos; FRY, Peter. A descoberta do Cafundó: alianças e conflitos no cenário da cultura negra no Brasil. *Religião e Sociedade,* São Paulo, n.8, p.45-52, 1982.

VOGT, Carlos; FRY, Peter. Ditos e feitos da *falange* africana do Cafundó e da *calunga* de Patrocínio (ou de como fazer falando). *Revista de Antropologia*, São Paulo, n.26, p.65-92, 1983.

VOGT, Carlos; FRY, Peter. Rios de cristal: contos e desencontros de línguas africanas no Brasil. *Cadernos de Estudos Linguísticos,* Campinas, n.8, p.109-128, 1985.

WEINREICH, U.; LABOV, W.; HERZOG, M. Empirical foundations for theory of language change. In: LEHMANN, P.; MALKIEL, Y. (Ed.). *Directions for Historical Linguistics*. Austin: University of Texas Press, 1968. p.95-188.

WEINREICH, U.; LABOV, W.; HERZOG, M. *Fundamentos empíricos para uma teoria da mudança linguística*. Trad. Marcos Bagno, revisão técnica Carlos Alberto Faraco. São Paulo: Parábola, 2006 [1968].

WHINNOM, Keith. Linguistic hybridization and the "special case" of pidgins and creoles. In: HYMES, D. (Ed.). *Pidginization and creolization of languages*. Cambridge: Cambridge University Press, 1971. p.91-115.

WILLIAMSON, Kay; BLENCH, Roger. Niger-Congo. In: HEINE, B.; NURSE, D. (Ed.). *African Languages*: an introduction. Cambridge, UK: Cambridge University Press, 2000. p.11-42.

WILSON, Warren A. A. *The Creole of Guiné*. Johannesburg: Witwatersrand University Press, 1962.

WINFORD, Don. "Intermediate" creoles and degrees of change in Creole formation: the case of Bajan. In: NEUMANN-HOLZSCHUH, I.; SCHNEIDER, E. W. (Ed.). *Degrees of restructuring in Creole languages*. Amsterdam: John Benjamins, 2000. p.215-274.

WINFORD, Don. *An introduction to contact linguistics*. Oxford: Blackwell, 2003.

WOOLFORD, Ellen. Bilingual code-switching and syntactic theory. *Linguistic Inquiry*, Massachusetts, v.14, n.1, p.520-536, 1983.

ZILLES, Ana; MAYA, Leonardo; SILVA, Karine. A concordância verbal com a primeira pessoa do plural na fala de Panambi e Porto Alegre, RS. *Organon*, Porto Alegre, v.14, n.28-29, p.195-220, 2000.

ZIMMERMANN, Klaus. O português não-padrão falado no Brasil: a tese da variedade pós-crioula. In: ______ (Ed.). *Lenguas criollas de base lexical española y portuguesa*. Madri: Iberoamericana, 1999. p.441-476.

ZUBIZARRETA, M. L. *Prosody, focus and word order*. Cambridge: The MIT Press, 1998.

ZUBIZARRETA, M. L. Las funciones informativas: tema y foco. In: BOSQUE, I.; DEMONTE, V. (Org.). *Gramática descriptiva de la lengua española*. v.3. Madri: Espasa Calpe, 1999. p.4215-4244.

Colofão

Formato	18 x 25 cm
Tipologia	IowanOldSt BT 10/14,5
Papel	Polén 80 g/m² (miolo) Cartão Supremo 250 g/m² (capa)
Impressão	Setor de Reprografia da EDUFBA
Capa e Acabamento	Bigraf
Tiragem	600